

Mestrado em Gestão de Informação

Master Program in Information Management

**A aplicação de self-organizing maps na
caracterização do universo geográfico dos
serviços de correio.**

Fernando António Martins Matos

Trabalho de projeto apresentado como requisito parcial
para obtenção do grau de Mestre em Gestão de
Informação

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

A APLICAÇÃO DE SELF-ORGANIZING MAPS NA CARACTERIZAÇÃO DO UNIVERSO GEOGRÁFICO DOS SERVIÇOS DE CORREIO.

por

Fernando António Martins Matos

Trabalho de projeto apresentado como requisito parcial para a obtenção do grau de Mestre em Gestão de Informação, Especialização em Gestão do Conhecimento e Business Intelligence.

Orientador: Professor Doutor Roberto Henriques

maio de 2017

RESUMO

Com a era do digital e do correio eletrónico, a utilização do correio físico tem vindo a decrescer e a cair em desuso. No entanto, o correio publicitário tem vindo a aumentar, verificando-se ainda que os produtos e serviços disponibilizados pelos CTT - Correios de Portugal, S.A. (CTT) são cada vez mais diversificados. Também o e-commerce, na vertente Business to Consumer (B2C), assume cada vez maior expressão no negócio dos CTT com a entrega das encomendas no domicílio do cliente. Com este trabalho pretende-se obter uma caracterização e segmentação da população cliente dos CTT. Partindo dos dados respeitantes às Loja Postal (Lj) e Centro de Distribuição Postal (CDP) dos CTT, foi efetuada uma segmentação inicial de Lj e CDP de acordo com as suas características. Posteriormente, foi possível criar conjuntos de freguesias associadas a estes mesmos segmentos, com o objetivo final de efetuar uma caracterização sociodemográfica utilizando os dados dos últimos censos. Visto que as variáveis utilizadas para construir estes segmentos (clusters), eram em grande número e com diferentes ordens de grandeza, foi necessário utilizar uma técnica de redes neuronais, mapas auto-organizáveis (Self Organizing-Maps, SOM), que permitiu reduzir a dimensionalidade de entrada a duas dimensões de saída. O resultado foi a identificação de quatro clusters de CDP e quatro de Lj, onde é notória a diferença entre os clusters com características urbanas e os outros com características rurais, bem como da sua especificidade quanto ao tipo de produtos e serviços utilizados. Na análise dos dados sociodemográficos foi possível distinguir os clusters com freguesias urbanas e rurais, essencialmente através das variáveis número de habitantes, alojamentos e pensionista/reformados, com valores médios muito diferentes, e dos níveis de educação, que variavam maioritariamente entre o 1º ciclo ou o ensino superior. A dispersão dos clusters de CDP e Lj no país, de uma forma geral, é semelhante.

PALAVRAS-CHAVE

Correios

Dados Sociodemográficos

Freguesias

Mapas Auto-organizáveis

População

ABSTRACT

The use of physical mail has been decreasing and falling into disuse as a result of the digital era and the electronic mail. On the other hand, advertising mail are increasing and so are the range of products and services rendered available by CTT- Correios de Portugal, S.A. (CTT), and in CTT business, increasing relevance is being given to e-commerce particularly in a Business to Consumer (B2C) scenario due to household package delivery. This work aims at obtaining a characterization and segmentation of CTT's customers. An initial segmentation of Post Offices (PO) and Postal Delivery Centers (CDP) was made in accordance with their features and taking in consideration the number of PO and CDP in operation. Subsequently, it was possible to create a group of parishes associated to these segments in order to make a social demographic characterization using the last census data. As variables used to build these segments (clusters) were many with several orders of magnitude, it was necessary to use a neuronal network technique, the self-organizing maps (SOM), which allowed entry resizing to two output variables. It resulted in the identification of four clusters of CDP and four PO, where the difference between urban and rural clusters is glaring and so is the specificity of products requested and services used. As a result of the analysis of social-demographic data it was possible to recognize which clusters are urban or rural, specifically by the variables "number of inhabitants", "lodgings" and "pensioners", with very different average values and education levels, ranging from primary school to university. CDP and PO clusters dispersion are much similar.

KEYWORDS

Mail

Parishes

Population

Self-organizing Map

Sociodemographic Data

Índice

1.	INTRODUÇÃO	1
2.	OBJETIVOS	2
3.	METODOLOGIA E DADOS	3
3.1.	DADOS DO INE	3
3.2.	DADOS DOS CTT	4
3.3.	ANÁLISE SOCIODEMOGRÁFICA	6
3.4.	ANÁLISE DE CLUSTERS	7
3.4.1.	SOM	7
4.	RESULTADOS	9
4.1.	SEGMENTAÇÃO DOS CDP	9
4.2.	SEGMENTAÇÃO DAS LJ	15
4.3.	ANÁLISE SOCIODEMOGRÁFICA	21
4.3.1.	POR CLUSTER DE CDP	21
4.3.2.	POR CLUSTER DE LJ	25
5.	CONCLUSÕES	27
6.	RECOMENDAÇÕES PARA TRABALHOS FUTUROS	28
7.	BIBLIOGRAFIA	29
8.	ANEXOS	31
8.1.	ANEXO 1	31
8.2.	ANEXO 2	32
8.3.	ANEXO 3	38

Índice de Imagens

Figura 1- Ilustração som, espaço de entrada e espaço de saída, (FONTE: Henriques, 2014)....	8
Figura 2 - Exemplo Matriz U (fonte: https://users.ics.aalto.fi/jhollmen/dippa/node24.html) .	8
Figura 3 - Matriz U dos CDP com labels	10
Figura 4 - Lista de variáveis utilizadas para o cluster 1 CDP.....	11
Figura 5 - Lista de variáveis utilizadas para o cluster 2 CDP.....	12
Figura 6 - Lista de variáveis utilizadas para o cluster 3 CDP.....	13
Figura 7 - Lista de variáveis utilizadas para o cluster 4 CDP.....	13
Figura 8 - Matriz U com clusters CDP	14
Figura 9 - Matriz U das Lj com labels.....	15
Figura 10 - Lista de variáveis utilizadas para o cluster 1 Lj	16
Figura 11 - Lista de variáveis utilizadas para o cluster 3 Lj	17
Figura 12 - Lista de variáveis utilizadas para o cluster 4 Lj	17
Figura 13 - Matriz U com clusters Lj.....	18
Figura 14 - Mapa de Portugal por Concelhos, clusters CDP	19
Figura 15 - Mapa de Portugal por Concelhos, clusters Lj.....	20
Figura 16 - Dados sociodemográficos do cluster 1 CDP.....	21
Figura 17 - Dados sociodemográficos do cluster 2 CDP.....	22
Figura 18 - Dados sociodemográficos do cluster 3 CDP	23
Figura 19 - Dados sociodemográficos do cluster 4 CDP	24
Figura 20 - Dados sociodemográficos do cluster 1 Lj.....	25
Figura 21 - Dados sociodemográficos do cluster 2 Lj.....	25
Figura 22 - dados sociodemográficos do cluster 3 Lj	26
Figura 23 - Dados sociodemográficos do cluster 4 Lj.....	27

Índice de Tabelas

Tabela 1 - Lista de variáveis utilizadas dos censos.....	4
Tabela 2 - Lista de variáveis utilizadas para as Lj	5
Tabela 3 - Lista de variáveis utilizadas para os CDP	5

LISTA DE SIGLAS E ABREVIATURAS

CDP	Centro de distribuição postal
CRT	Carteiro
CTT	CTT - Correios de Portugal, S.A.
e-mail	Correio eletrónico
e-commerce	Comércio eletrónico
INE	Instituto Nacional de Estatística, IP
LJ	Loja postal
SOM	Mapas auto-organizáveis

1. Introdução

A utilização do correio físico está a decrescer e a cair em desuso, como se constatou no último relatório internacional (“Universal Postal Union – 2013 Results,”2013), e também em Portugal (“ANACOM - Serviços Postais - 4.º trimestre de 2014,”2014). Para tal tem contribuído a massificação da utilização do correio eletrónico (e-mail). Não obstante, há ainda correspondência que por imposição legal tem que ser em formato físico, como seja o correio registado, e clientes que por vontade própria ou por limitação de meios, continuam a enviar e receber correspondência em formato físico. Pelas situações referidas e pelo restante tipo de correio que circula na rede dos CTT - Correios de Portugal, S.A. (CTT), correio empresarial, publicitário e particular, o correio físico continuará a ser uma realidade.

Os CTT são uma empresa com dispersão a nível de Portugal Continental e Ilhas, e detentores de uma vasta rede de Centro de Distribuição Postal (CDP) e Loja Postal (Lj), 252 CDP e 606 Lj, à data deste trabalho. As Lj, antes chamadas Estações de Correio (EC), são o local onde os clientes se deslocam sempre que pretendem utilizar um serviço ou adquirir um produto dos CTT. Os CDP são o local de onde saem os carteiros (CRT) para os respetivos giros. O giro é o trajeto que um CRT efetua para entrega de correspondência no domicílio dos clientes, podendo também incluir a recolha de correio dos marcos de correio ou ainda a venda de produtos, como por exemplo selos postais. Cada Lj relaciona-se com um CDP e os giros de um CDP delimitam a sua área de influência, que pode abranger uma ou mais freguesias. Desta forma é possível criar a associação entre freguesias e CDP, e freguesias e Lj.

Para a caracterização do universo geográfico dos serviços de correio foi necessário realizar uma segmentação de CDP e Lj. Partindo dos segmentos criados foi possível identificar as freguesias associadas a cada segmento de Lj e CDP, e assim efetuar uma análise sociodemográfica sobre a população residente nas freguesias de cada segmento. Os dados dos CTT referentes ao ano de 2015 foram utilizados para a segmentação, e para a análise sociodemográfica os disponibilizados pelo INE, resultantes dos últimos censos de 2011.

Para criar segmentos de CDP e Lj foi efetuada uma revisão de literatura sobre estudos na área, de forma a identificar as técnicas utilizadas neste tipo de problemas. A literatura encontrada refere-se à segmentação de clientes do sector energético, de forma a classificar e segmentar grupos de clientes, com o objetivo de tirar mais valor do negócio e disponibilizar melhores serviços aos clientes. Neste trabalho podemos efetuar o paralelismo de que as Lj e os CDP são clientes com características específicas no negócio dos CTT. Na literatura

analisada encontraram-se várias técnicas (Chicco, Napoli, & Piglion, 2006) mas as que se destacaram foram o *k*-means e os SOM - Self Organizing-Maps (SOM), ambas são tipos de redes-neuronais com aprendizagem não supervisionada, e utilizadas para criar segmentações. Foi analisada literatura onde estas técnicas foram utilizadas em conjunto, ou apenas o *k*-means mas com alterações ao algoritmo original (López et al., 2011)(Benítez, Quijano, Díez, & Delgado, 2014). Nas metodologias utilizando as duas técnicas em conjunto, enquanto numa, os SOM tomaram o papel de redutor de dimensionalidade e depois foi utilizado *k*-means de forma a criar os clusters (Li & Lin, 2008), na outra, os SOM foram utilizados de forma a agrupar os dados de input pelas suas características, e o *k*-means utilizado para a definição do número de clusters para o resultado final (Sánchez, Espinós, Sarrión, López, & Burgos, 2009)(Verdú, García, Senabre, Marín, & Franco, 2006).

Em todos os documentos analisados foi notório que os SOM constituem uma técnica com bons resultados ao nível da segmentação e redução da dimensionalidade, sendo os seus resultados facilmente analisados visualmente, permitindo a identificação de padrões, relações ou segmentos.

Neste trabalho os SOM foram utilizados para efetuar a segmentação, dado o grande número de variáveis, e algumas delas com ordens de grandeza diferentes. Na definição do número de segmentos, foi utilizado a matriz U e o plano de componentes que se desenvolverá mais à frente. Desta forma foi possível verificar o peso de cada variável na definição de cada segmento, e a respetiva dispersão no universo dos CDP e Lj.

2. Objetivos

Na elaboração deste trabalho foram estabelecidos três objetivos: dois específicos, segmentação e caracterização dos CDP e segmentação das Lj por hábitos de consumo de produtos e serviços pelos clientes, e um final que agregará os dois anteriores, com vista à caracterização do universo geográfico dos serviços dos CTT. Com o primeiro objetivo pretende-se criar segmentos (clusters) de CDP com características semelhantes entre si, de forma a identificar padrões de tipo de correio utilizado, bem como as características da área servida por estes CDP quando à sua ruralidade. Com estes clusters obtém-se uma lista de freguesias para posterior análise sociodemográfica. No segundo objetivo e partindo dos dados sobre vendas nas Lj, pretende-se criar clusters de acordo com os hábitos de consumo de produtos e serviços por parte dos clientes. Os clusters obtidos permitem organizar uma lista

de freguesias para posterior análise sociodemográfica. Com o objetivo final será feita uma análise sociodemográfica a cada lista de freguesias, ou seja, a cada cluster de Lj e CDP.

Por fim serão elaborados dois mapas de Portugal com o resultado da segmentação das Lj e CDP, onde será possível verificar a dispersão geográfica ao nível dos concelhos.

3. Metodologia e Dados

Definidos os objetivos deste trabalho, foi necessário escolher as variáveis que melhor caracterizam as Lj e CDP. Neste passo efetuaram-se várias simulações utilizando os SOM, com diferentes combinações de variáveis, e diferentes valores para os parâmetros utilizados nos SOM, para que a escolha das variáveis que respondessem aos objetivos fixados fosse a mais correta. Esta seleção foi feita por tentativa e erro, análise de resultados bem como a análise das medidas de avaliação resultantes da aplicação do SOM, erro de quantização e erro topológico. Por fim, o resultado de cada segmentação e respetiva caracterização foi analisado de forma a perceber se os resultados obtidos se apresentavam coerentes com os objetivos traçados e com a realidade de negócio dos CTT.

3.1. Dados do INE

Os Recenseamentos Gerais da População e da Habitação (Censos) elaborados pelo Instituto Nacional de Estatística (INE) têm como principal objetivo a caracterização da população do país. Apesar de os censos serem uma análise estatística com diversas variáveis de estudo, para este trabalho só foram utilizadas 20 variáveis recolhidas nos últimos censos com data em 2011 (fonte: http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=censos_quadros).

Nome variável	Descrição	Nome variável	Descrição
Homens	Quantidade de habitantes do sexo masculino	3º Ciclo	Quantidade de habitantes com o 3º ciclo
Mulheres	Quantidade de habitantes do sexo feminino	Secundário	Quantidade de habitantes com ensino secundário
0-14	Quantidade de habitantes na faixa etária dos 0 aos 14 anos	Pós-secundário	Quantidade de habitantes com ensino pós-secundário
15-24	Quantidade de habitantes na faixa etária dos 15 aos 24 anos	Superior	Quantidade de habitantes com ensino secundário
25-64	Quantidade de habitantes na faixa etária dos 25 aos 64 anos	Num de desempregados	Quantidade de habitantes desempregados
65/+	Quantidade de habitantes na faixa etária superior a 65 anos	Total de empregados	Quantidade de habitantes empregados
Alojamentos	Quantidade de alojamentos	Pensionistas/reformados	Quantidade de habitantes na situação de Pensionistas ou reformados
Nenhuma	Quantidade de habitantes sem escolaridade	Primário	Quantidades de habitantes trabalhadores no sector primário
1º Ciclo	Quantidade de habitantes com o 1º ciclo	Secundário	Quantidades de habitantes trabalhadores no sector secundário
2º Ciclo	Quantidade de habitantes com o 2º ciclo	Terciário	Quantidades de habitantes trabalhadores no sector terciário

TABELA 1 - LISTA DE VARIÁVEIS UTILIZADAS DOS CENSOS

Aquando da realização dos últimos censos ainda não vigorava a atual Carta Administrativa Oficial de Portugal (CAOP). Esta reorganização administrativa das freguesias ocorreu em 2013, e originou a redução do número de freguesias de 4260 para 3092. Para a utilização dos dados dos censos de 2011 com os dados dos CTT de 2015, já com base na nova organização administrativa das freguesias de 2013, foi necessária a agregação dos respetivos dados, para as freguesias que foram unidas numa nova freguesia e, para as freguesias que se extinguíram dando origem a mais do que uma, proceder à divisão dos dados nas nova freguesias. Por ex. a antiga freguesia de Santa Maria dos Olivais foi distribuída pelas freguesias do Lumiar, Olivais, Parque das Nações e Santa Clara, tendo os respetivos dados sido divididos por quatro. A necessidade desta divisão só ocorreu em 12 freguesias, todas as outras ou foram agregadas ou mantiveram-se inalteradas.

3.2. Dados dos CTT

Os CTT possuem uma vasta rede de Lj e de CDP, no continente e ilhas, estabelecimentos estes que produzem uma grande quantidade de dados diários, que são agregados para produção de indicadores mensais e anuais. Foi com base nos dados acumulados de 2015, utilizados para KPI, que se seleccionou uma lista de variáveis suficientemente caracterizadores de forma a criar clusters de Lj e CDP. Para as Lj foram escolhidas 14 variáveis.

Nome variável	Descrição
Receita Ocasional	Valor da receita aceite na Lj de clientes ocasionais
Receita Contratual	Valor da receita aceite na Lj de clientes contratuais
Correspondências	Valor da receita de correio entregue na Lj por clientes ocasionais
Direct mail	Valor da receita de Direct mail entregue na Lj por clientes ocasionais
Colecionismos	Valor de receita obtido pela venda de produtos de colecionismos (ex.: selos de coleção) por clientes ocasionais
Outros Prod e serviços	Valor de receita obtido pela venda de outros produtos e serviços (ex.: embalagens) por clientes ocasionais
Serv Financeiros	Valor de receita obtido pela venda de produtos de serviços financeiros (ex.: certificados de aforro) por clientes ocasionais
Retalho	Valor de receita obtido pela venda de produtos de Retalho (ex.: livros) por clientes ocasionais
Subsidiárias	Valor de receita obtido pela venda de produtos ou serviços das empresas subsidiárias por clientes ocasionais
Clientes/DIA	Média de clientes dia
CTT Express	Valor de receita obtido pela venda de produtos da CTT Express por clientes ocasionais
Post Contact	Valor de receita obtido pela venda de produtos da Post Contact por clientes ocasionais
ETI	Número de colaboradores a tempo inteiro

TABELA 2 - LISTA DE VARIÁVEIS UTILIZADAS PARA AS LJ

Na lista de variáveis utilizadas importa salientar que a receita contratual advém de contratos celebrados com clientes. A receita ocasional provém dos clientes ocasionais na Lj e é a soma das seguintes variáveis: Correspondências, Direct Mail, Encomendas, Colecionismos, Outros Produtos e serviços, Serviços Financeiros, Retalho e Subsidiárias. A variável Clientes/dia é a média diária num ano. CTT Express e Post Contact são também tipos de produtos disponibilizados pelos CTT. Por fim as ETI representam o número de colaboradores a tempo inteiro.

Para os CDP foram escolhidas 22 variáveis que possibilitam a caracterização do CDP quanto ao tipo de correio distribuído bem como à ruralidade da área servida pelo CDP.

Nome variável	Descrição	Nome variável	Descrição
Dv	Quantidade de giros do tipo Diversos Viatura	Apartados	Percentagem de correio de um CDP entregue em apartados
Mistos	Quantidade de giros do tipo Mistos	Ent Diretas	Percentagem de correio de um CDP com entrega direta ao cliente
Rural	Quantidade de giros do tipo Rural	RSF	Percentagem de correio de um CDP entregue em RSF
Urbanos	Quantidade de giros do tipo Urbanos	EE	Percentagem de correio de um CDP com correio errado, pertencente a outro CDP
Terceiros	Quantidade de giros do tipo Terceiros	ETI	Número de colaboradores a tempo inteiro
CTT	Quantidade de viaturas propriedade dos CTT	Registos	Quantidade de correio do tipo de Registos
CRT	Quantidade de viaturas propriedade dos CRT	Encomendas	Quantidade de correio do tipo de Encomendas
Terceiros	Quantidade de viaturas propriedade de terceiros	EMS	Quantidade de correio do tipo de EMS (CTT Express)
Vendas CRT	Valor das vendas efetuadas por CRT	Correio não endereçado	Quantidade de correio do tipo de Correio não endereçado
CRT	Percentagem de correio de um CDP entregue pelos CRT	Registos Avisados (%)	Percentagem de registos avisados
Agenciados	Percentagem de correio de um CDP entregue por giros agenciados a terceiros	Correio Normal e Direct Mail	Quantidade de correio normal e direct mail

TABELA 3 - LISTA DE VARIÁVEIS UTILIZADAS PARA OS CDP

Nesta lista de variáveis salientamos o Tipos de itinerários (giros). Em cada CDP, é a quantidade de giros de cada Tipo que caracteriza a respetiva área de influência. As variáveis integradas no Tráfego são caracterizadoras do tipo de correio que é entregue pelos CRT. Os ETI representam o número de CRT. As vias de saída representam o meio como foi entregue o correio, pelos CRT, em apartados, por agenciados, entregas diretas (empresas), RSF, EE (correio mal endereçado). São ainda relevantes para a caracterização as Vendas CRT, valor em euros das vendas efetuadas pelos CRT durante o giro, e a informação dos proprietários das viaturas utilizadas no CDP.

3.3. Análise Sociodemográfica

A área geográfica onde cada CDP presta serviço não corresponde à divisão administrativa de Portugal, uma vez que os CDP trabalham por áreas que são definidas pelos CTT, tendo em conta determinadas variáveis, como o tráfego dos vários tipos de correio, número de pontos de distribuição (habitações), distâncias, relevo da região entre outras, que assim definem os vários giros de um CDP.

Deste modo, uma mesma freguesia pode ser servida por mais que um CDP e, em consequência, surgir em clusters diferentes quando esses CDP não tenham ficado no mesmo cluster.

Dentro de cada cluster os dados foram trabalhados de forma a não haver repetições de freguesias. Ainda assim, existem 37 freguesias repetidas, isto é, freguesias que constam em mais que um cluster por serem servidas por CDP que ficaram colocados em clusters diferentes. Estas repetições acontecem na sua maioria nos concelhos de Lisboa e Porto. Optei por manter estas repetições por não deturparem o resultado final.

Quanto às Lj, sabendo que só se relacionam com um CDP, ligação entre a Lj e as freguesias obtém-se através desse CDP. Contudo, como um CDP pode ter na sua área de influência mais que uma LJ, as freguesias associadas ao CDP podem surgir repetidas quando as LJ da sua área de influência tenham ficado em clusters diferentes. Nesta situação temos 1097 freguesias, que aparecem repetidas em clusters diferentes. Neste caso, os dados de cada freguesia foram divididos pelo número de repetições, de forma a não alterar a análise do resultado.

Por existirem freguesias que contêm na sua área geográfica mais que uma LJ, os dados referentes a essas freguesias não foram divididos. Nesta situação temos 112 freguesias em que os seus dados ficaram repetidos em clusters diferentes.

Após terem sido criadas estas listas de freguesias, por cluster de CDP e Lj, foi elaborado um quadro de estatística descritiva utilizando o Excel, de forma a efetuar uma análise sociodemográfica do universo geográfico de utilizadores dos serviços e produtos dos CTT. As medidas usadas de tendência central foram a média, mediana e moda, as de variabilidade foram o desvio padrão, variância, o valor máximo e mínimo.

Por fim estas freguesias foram agregadas nos respetivos concelhos de forma a criar um mapa de cluster de CDP e Lj, onde é possível analisar a dispersão dos clusters pelo país.

3.4. Análise de clusters

3.4.1. SOM

Os SOM são uma técnica baseada em redes neuronais utilizada na segmentação de dados que permite reduzir n dimensões de entrada a duas dimensões de saída, para mais fácil compreensão e visualização dos dados, entre outras finalidades (Kohonen, 2001). Os SOM consistem na colocação de uma matriz de neurónios no espaço de entrada dos dados a analisar, e de sucessivas iterações, designadas por aprendizagem, em que estes neurónios se irão aproximar dos seus dados mais próximos, até obter uma forma de representar os dados garantindo que as suas características e relações topológicas sejam mapeadas para duas dimensões no espaço de saída (Bação, Lobo, & Painho, 2005)(Henriques, Lobo, & Bacao, 2010)(Figura 4). Na visualização dos resultados da técnica SOM, existem dois métodos que são os mais utilizados: o dos planos de componentes (Kohonen, 2001) que consiste em colorir os neurónios com um gradiente de cor que representa o peso de cada variável, e as matrizes U (Vesanto, 1999)(Figura 5) que representam as distâncias entre neurónios no espaço de entrada. Para esta técnica também é utilizado um gradiente, mas de cinzentos, em que os mais claros representam os neurónios mais próximos (maior semelhança) e os mais escuros representam os com maior distância (menor semelhança).

Esta última técnica facilita a visualização de fronteiras entre grupos semelhantes, sendo as cores mais escuras as fronteiras.

Para a criação dos clusters, e para ser possível utilizar esta técnica, é necessário preparar os dados e configurar alguns parâmetros (Vesanto, Himberg, Alhoniemi, & Parhankangas, 2000). Nos SOM, se existirem dados com ordens de grandeza muito diferentes, estes podem deturpar os resultados. A fim de evitar esta situação, efetuou-se uma normalização dos dados utilizando o valor de variância 1, ou seja, todos os valores das variáveis foram reduzidos a um intervalo entre $[0,1]$. De seguida foi definida a estrutura da matriz como sendo hexagonal, o número de neurónios na matriz 20x20, o raio de vizinhança, a taxa de aprendizagem e o número de épocas durante a qual é feito a aprendizagem.

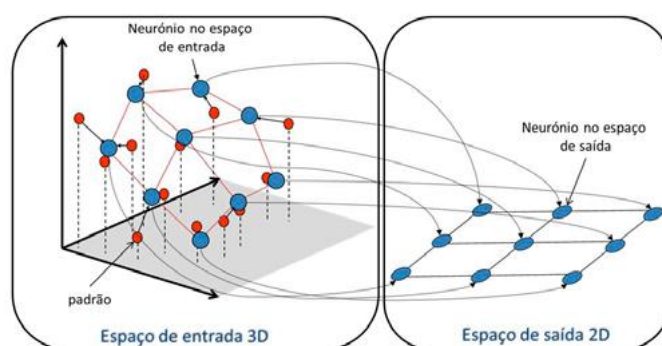


FIGURA 1- ILUSTRAÇÃO SOM, ESPAÇO DE ENTRADA E ESPAÇO DE SAÍDA, (FONTE: HENRIQUES, 2014)

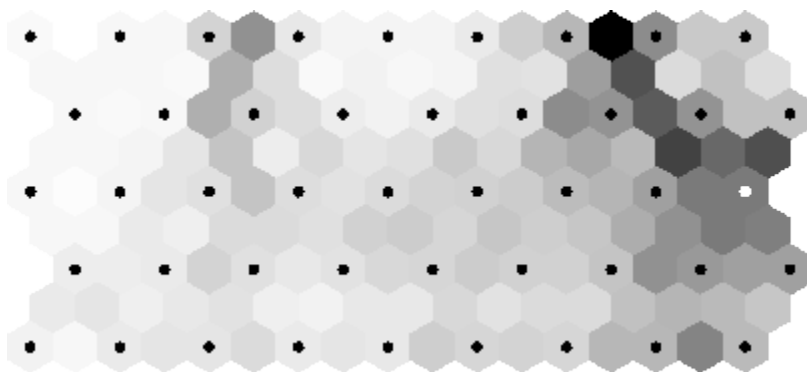


FIGURA 2 - EXEMPLO MATRIZ U (FONTE: [HTTPS://USERS.ICS.AALTO.FI/JHOLLMEN/DIPPA/NODE24.HTML](https://users.ics.aalto.fi/jhollmen/dippa/node24.html))

Para que o processo de aprendizagem e aproximação dos neurónios aos dados seja o mais fidedigno é necessário efetuar dois treinos, o primeiro que podemos chamar de dispersão, onde os parâmetros utilizados são elevados, o que faz com que os neurónios se dispersem pelos dados e abranjam a nuvem de dados de uma forma integral; o segundo onde os parâmetros são de valor reduzido, que se pode chamar de afinação, em que os neurónios se

vão deslocar menos e assim aproximar-se dos dados que estão mais perto de forma a representá-los de uma forma mais correta (Verdú et al., 2006). Não existe uma fórmula para calcular os parâmetros corretos, este é um processo empírico de teste, análise e repetição até produzir resultados coerentes com o problema.

Para este trabalho foi utilizado o programa MATLAB R2016 com a SOM toolbox 2.0 (fonte: <http://www.cis.hut.fi/somtoolbox/>)

4. Resultados

4.1. Segmentação dos CDP

Na análise de resultados obtidos pelo SOM começou-se pela matriz U, para se perceber como ficaram organizados os CDP, passando depois a analisar a contribuição que cada variável (plano de componentes) teve no resultado final. Este plano de componentes, que representa a variação da variável na matriz, mostra como essa variável ficou distribuída e o peso que teve na forma como os CDP se agruparam.

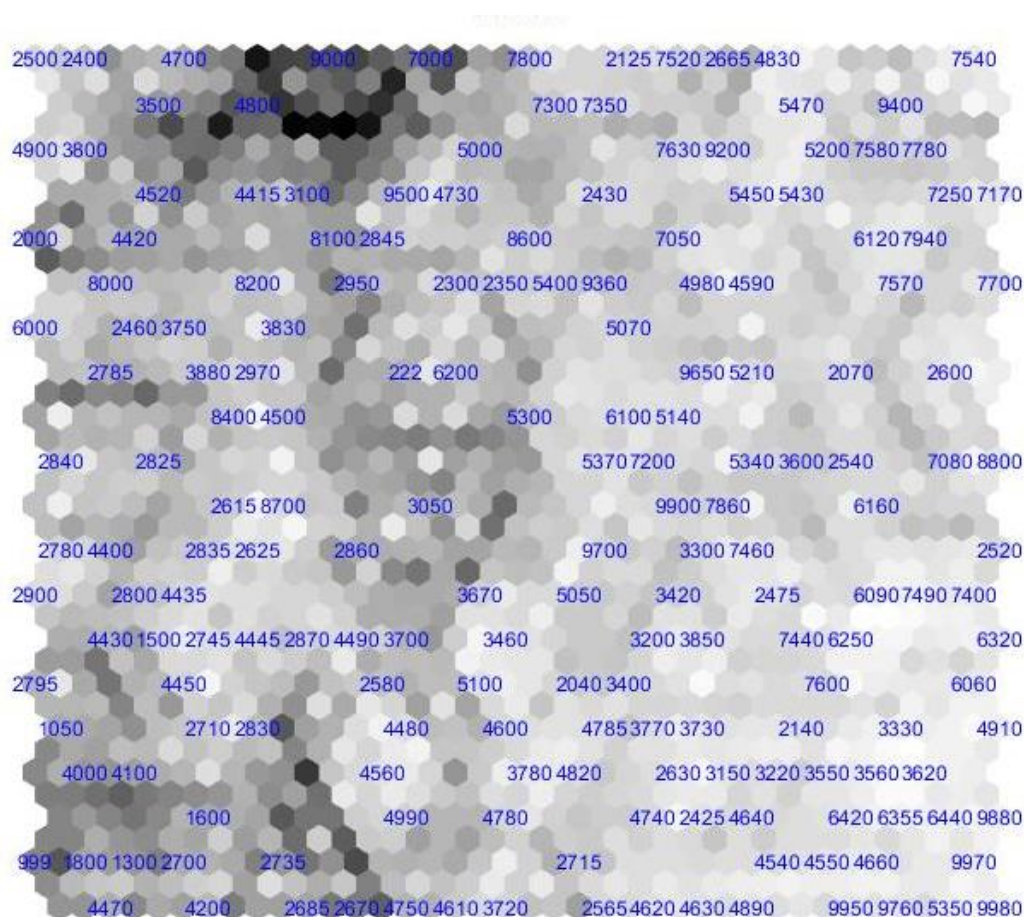


FIGURA 3 - MATRIZ U DOS CDP COM LABELS

Nesta matriz U é possível verificar as distâncias entre os neurónios e a localização de cada CDP na matriz, representado pelos seus 4 primeiros dígitos do código postal, sendo que esta localização pode integrar mais que um CDP (Anexo 1). No lado direito existe uma maior homogeneidade entre as distâncias, pelo que estes CDP terão características semelhantes. No lado esquerdo, pelo contrário, existe uma maior diversidade de distâncias.

Como o resultado da matriz U não permitiu definir clusters, foi também utilizado o plano de componentes, com uma escala de azuis para valores mais baixos e amarelos para valores elevados. Após a análise conjunta da matriz U com o plano de componentes foi possível a definição de 4 clusters.

Para o cluster 1 as variáveis com maior peso para a definição da segmentação foram:

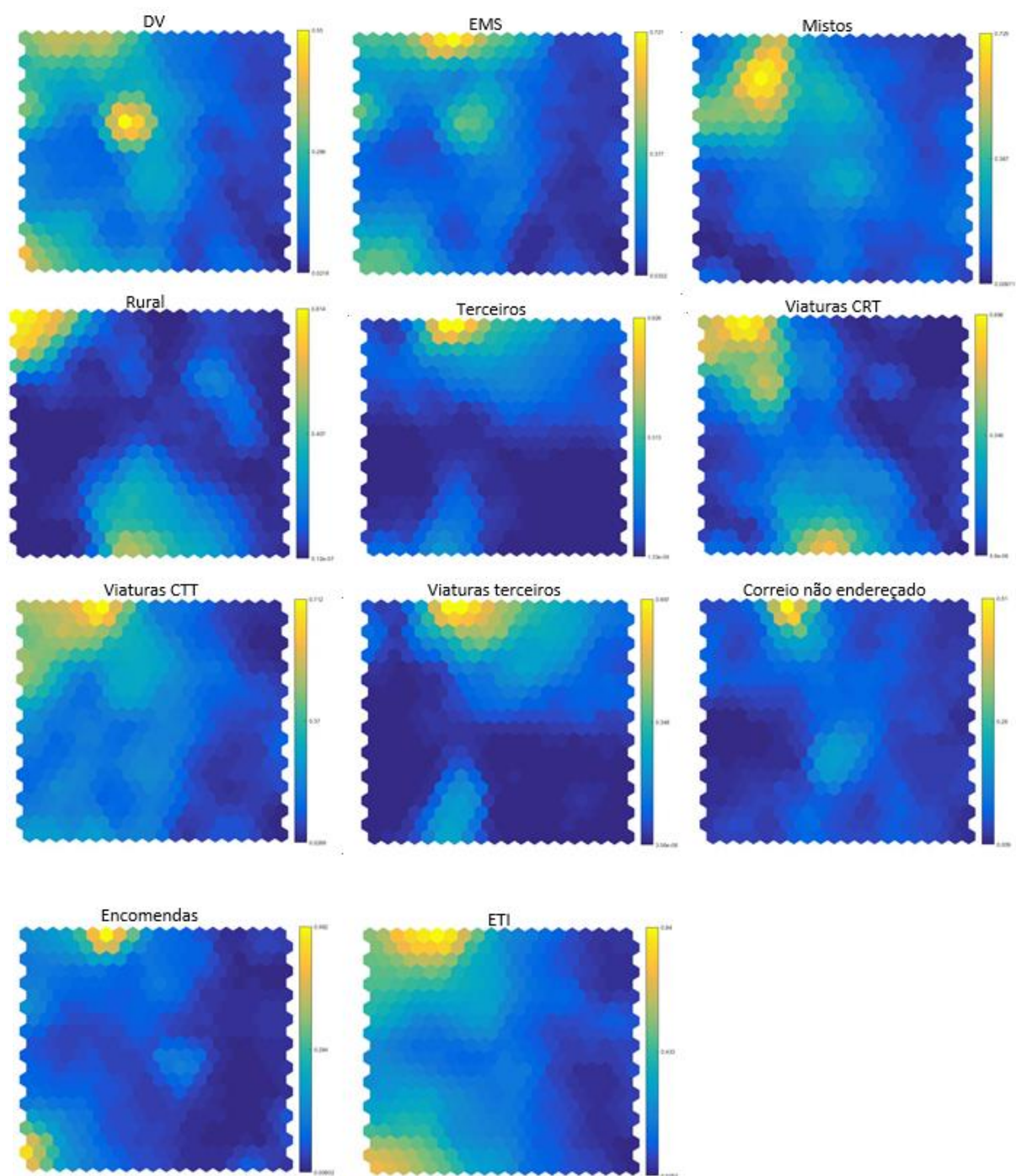


FIGURA 4 - LISTA DE VARIÁVEIS UTILIZADAS PARA O CLUSTER 1 CDP

Verifica-se que nestas 11 variáveis as agregações com valores mais elevados se situam maioritariamente na área do canto superior esquerdo. Sendo que este cluster se caracteriza por ter giros mistos, rurais e de terceiros, bem como viaturas dos CTT e CRT. Podemos concluir que aqui ficaram CDP situados em zonas mais rurais, onde os giros são feitos maioritariamente em viatura e não apeados. Os giros de terceiros, bem como viaturas de

terceiros estão correlacionados por se tratarem de entidades externas aos CTT, sendo visível esta correlação por terem uma dispersão muito semelhante. Este cluster contém 18 CDP.

Na seleção do cluster 2 as variáveis mais significativas e com maior peso foram as seguintes:

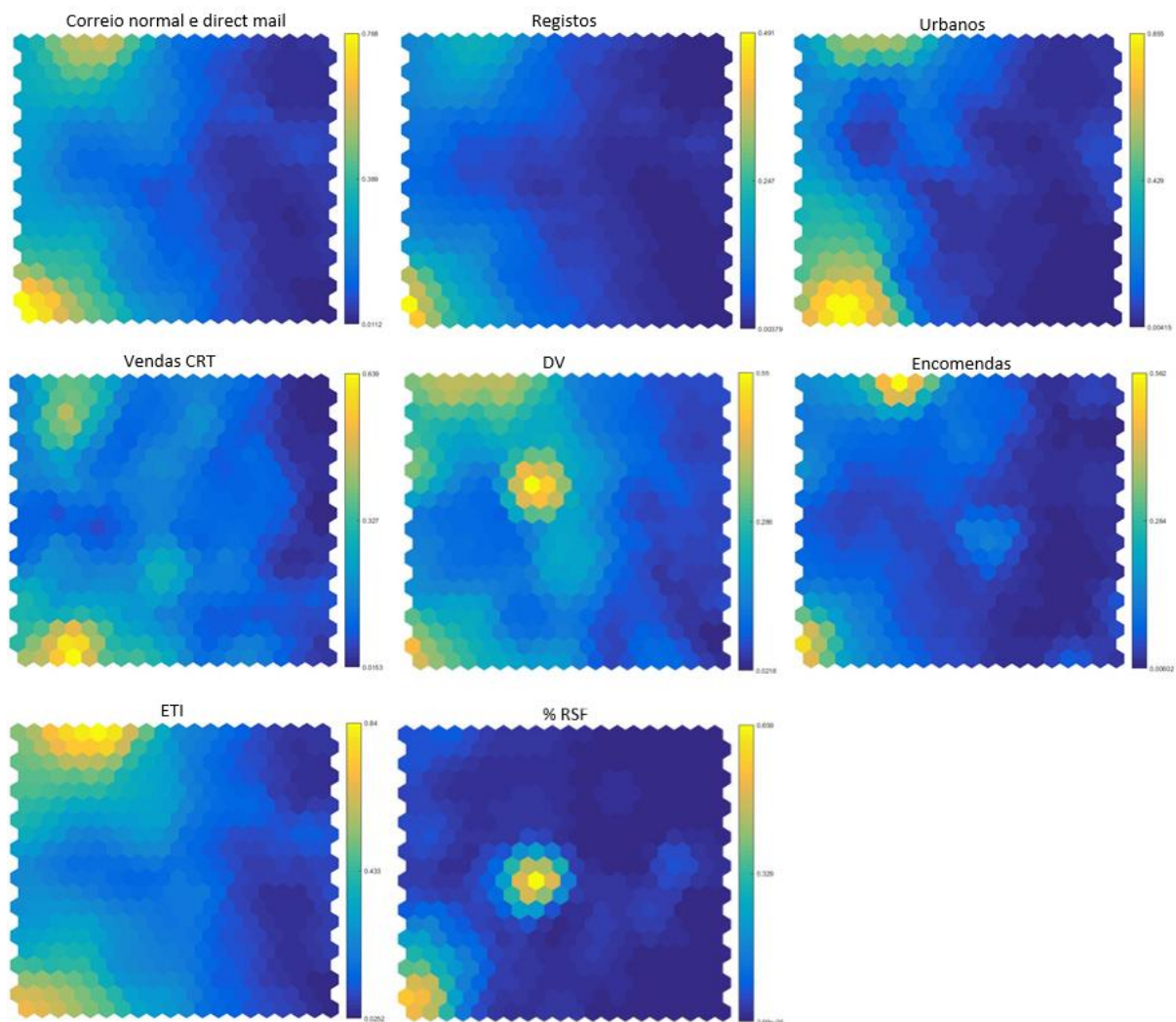


FIGURA 5 - LISTA DE VARIÁVEIS UTILIZADAS PARA O CLUSTER 2 CDP

Nestas variáveis podemos verificar que as maiores agregações se situam no canto inferior esquerdo. Verifica-se também que existem algumas variáveis que além de terem uma zona com maior concentração, têm outros pontos onde se situaram valores elevados. Esta situação demonstra que os CDP em causa ficaram alojados em outra zona devido à influência de outras variáveis. Este cluster fica caracterizado pelos giros urbanos, correio normal e direct mail, vendas CRT, registos e % RSF, aqui residindo os CDP das grandes cidades e associados a empresas. Este cluster contém 13 CDP.

Na criação do cluster 3 as variáveis com maior influência foram:

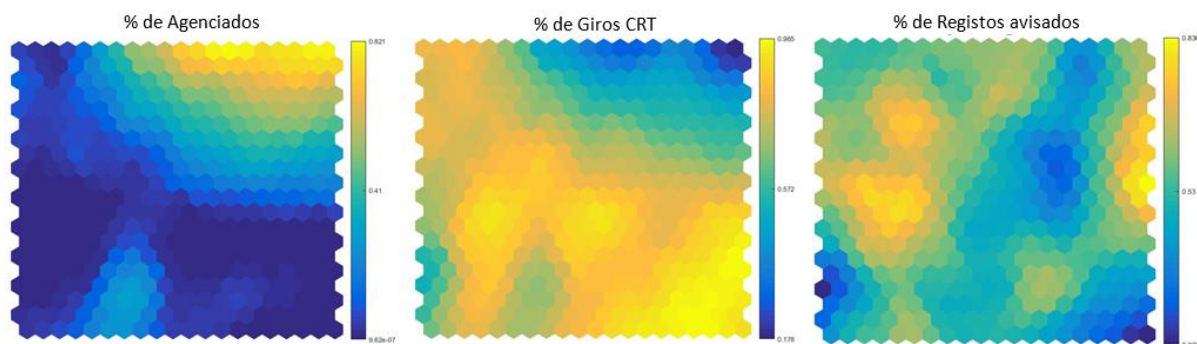


FIGURA 6 - LISTA DE VARIÁVEIS UTILIZADAS PARA O CLUSTER 3 CDP

Neste cluster ficaram os CDP com maior quantidade de correio entregue por agenciados, Giros CRT e com Registos avisados, estando as maiores agregações concentradas no lado direito.

É possível constatar que as variáveis % de giros CRT e % Registos avisados têm uma grande dispersão, significando que não existe um grupo pequeno de CDP que se destaque com valores elevados, são variáveis com grande homogeneidade de valores para todos os CDP. Este cluster contém 157 CDP.

Na criação do cluster 4 as variáveis com maior influência foram:

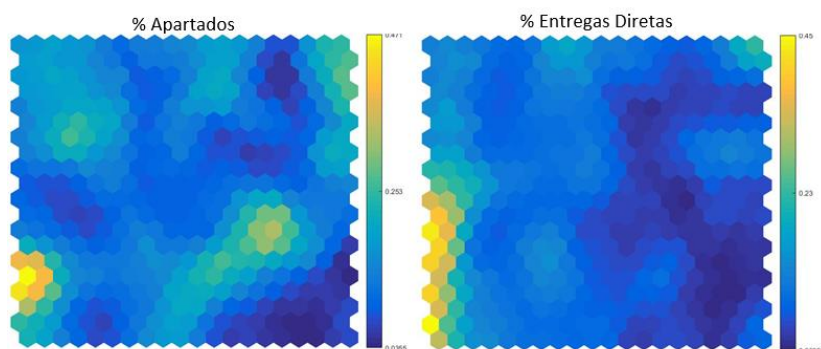


FIGURA 7 - LISTA DE VARIÁVEIS UTILIZADAS PARA O CLUSTER 4 CDP

O cluster 4 ficou com a área do lado esquerdo central. As variáveis com maior peso para este cluster foram % Apartados e % Entregas diretas. As variáveis % giros CRT e % de registros têm também aqui uma forte representação. Este cluster contém 64 CDP.

Os clusters dos CDP ficaram assim definidos.

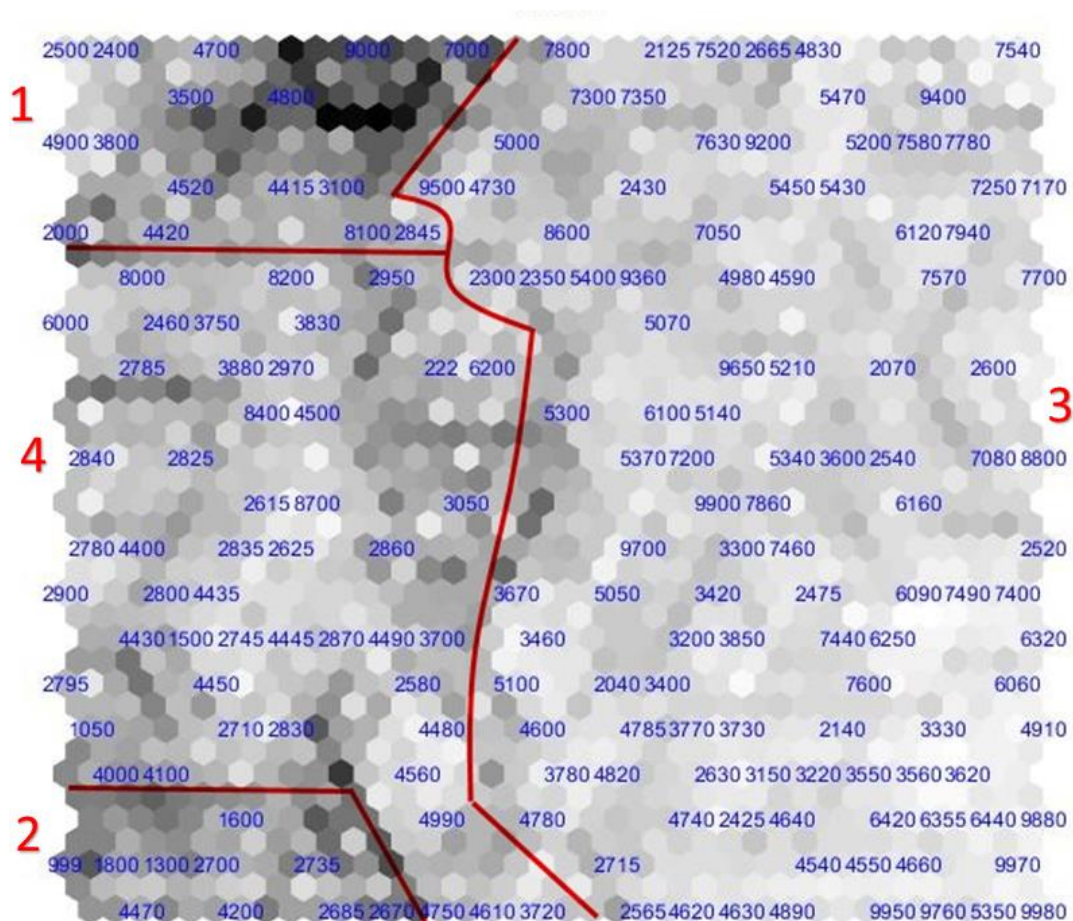


FIGURA 8 - MATRIZ U COM CLUSTERS CDP

4.2. Segmentação das Lj

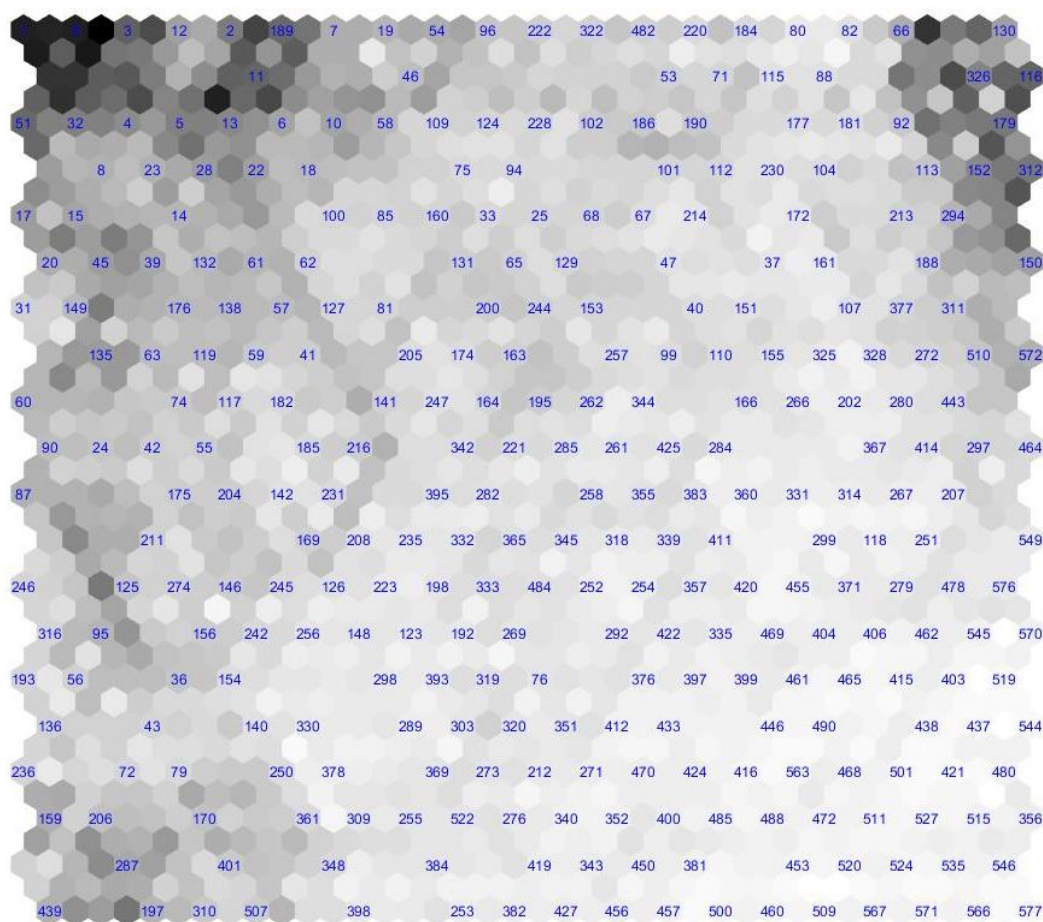


FIGURA 9 - MATRIZ U DAS Lj COM LABELS

De forma a ser visível graficamente, os nomes das Lj foram substituídos por um ID (Anexo 2), sendo que, tal como nos CDP, cada localização pode integrar mais que uma Lj (Anexo 3).

É possível verificar uma grande homogeneidade em grande parte da matriz, exceto no canto superior esquerdo e direito. Como o resultado da matriz U também não permitiu definir clusters, foi novamente utilizado o plano de componentes, aqui com uma escala de cinzentos, sendo o valor mais alto o preto e o mais baixo o branco.

Após a análise conjunta da matriz U com o plano de componentes foram definidos 4 clusters.

Para a definição do cluster 1 as variáveis com maior peso foram:

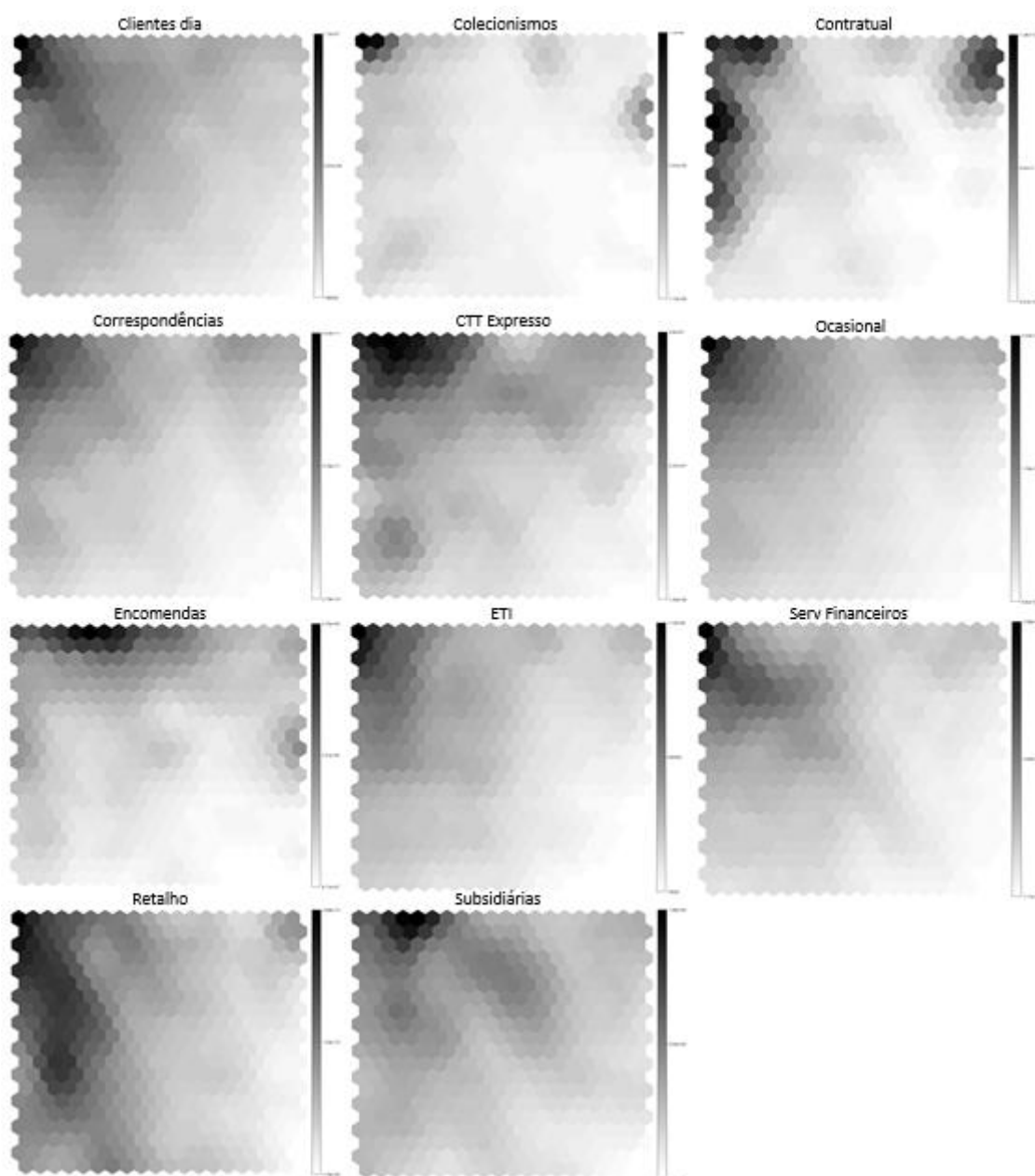


FIGURA 10 - LISTA DE VARIÁVEIS UTILIZADAS PARA O CLUSTER 1 Lj

Neste cluster ficaram as variáveis com maior concentração de valores elevados no canto superior esquerdo. Aqui estão representadas Lj com um grande número de Clientes/dia e com produtos específicos como o Colecionismo, Serviços Financeiros, CTT Expresso e Retalho. Vem igualmente representada a variável Contratual e Ocasional, que diz respeito ao tipo de clientes na Lj. Destes aspetos pode concluir-se que neste cluster estão Lj localizadas em grandes cidades, e que têm também empresas como clientes. Este cluster contém 128 Lj.

Na seleção do cluster 2 não houve nenhuma variável que tivesse grande destaque ou peso, todas têm valores homogêneos, desta forma foi criado o cluster 2 que se caracteriza por ter valores médios/baixos em todas as variáveis, correspondendo a Lj situadas em cidades e vilas mais pequenas, com menor tráfego e clientes. Este cluster contém 423 Lj.

As variáveis que tiveram mais peso na criação do cluster 3 foram as seguintes:

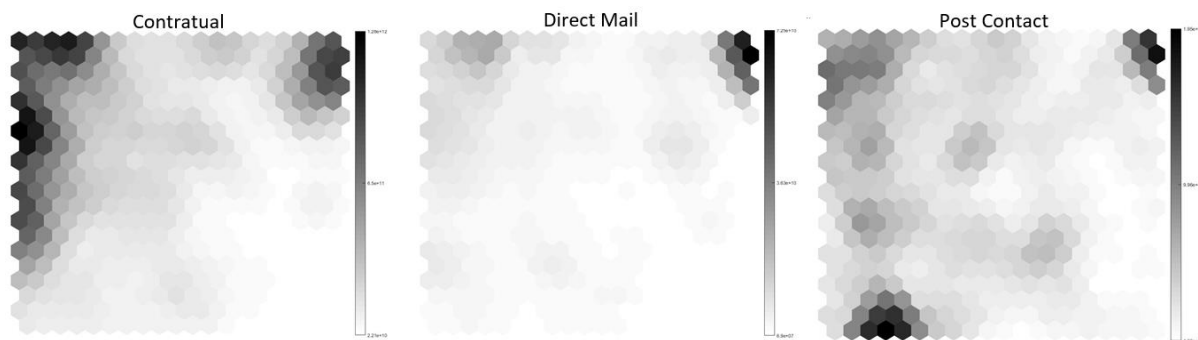


FIGURA 11 - LISTA DE VARIÁVEIS UTILIZADAS PARA O CLUSTER 3 Lj

Estas variáveis têm como padrão valores mais elevados no canto superior direito. O cluster 3 fica caracterizado por Direct mail, Post contact e Contratual. Este tipo de produtos é mais utilizado por empresas. Este cluster contém 36 Lj.

No cluster 4 as variáveis com mais peso foram as seguintes:

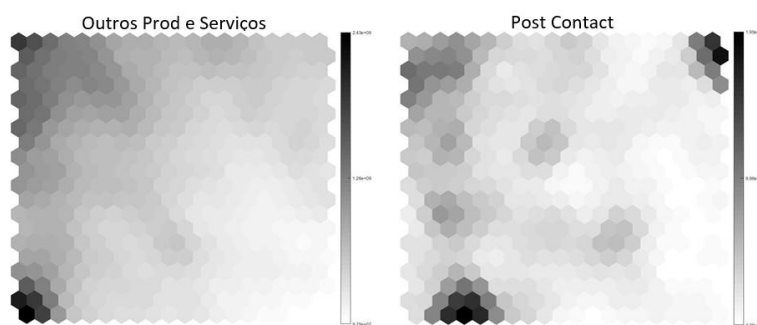


FIGURA 12 - LISTA DE VARIÁVEIS UTILIZADAS PARA O CLUSTER 4 Lj

No cluster 4 as variáveis com mais peso foram Post Contact e outros Produtos e serviços, que têm os valores mais elevados no canto inferior esquerdo, constituindo a característica deste cluster. Este cluster é o mais pequeno, não ficando claro nenhum tipo de padrão para as Lj. Este cluster contém 19 Lj.

Os clusters das Lj ficaram assim definidos.

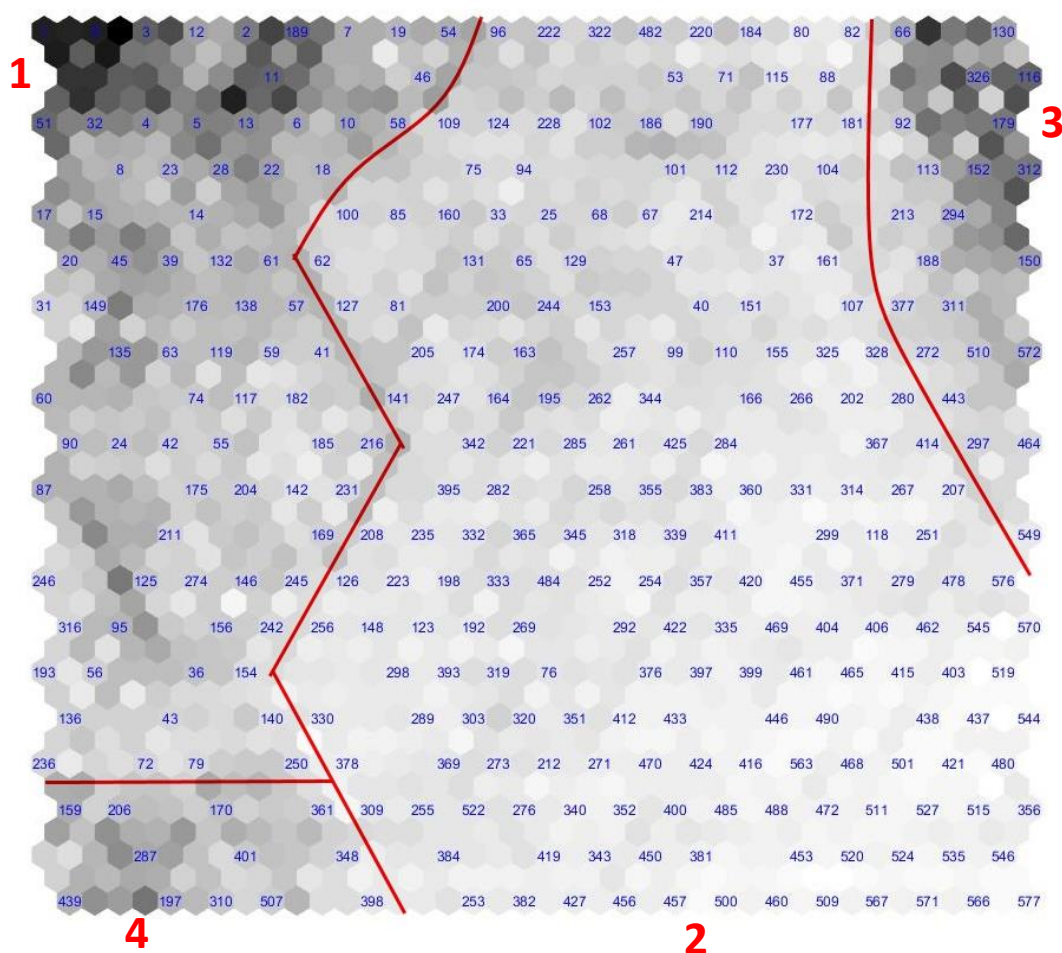


FIGURA 13 - MATRIZ U COM CLUSTERS LJ

Como foi referido na metodologia, depois de criados os clusters, é possível saber as freguesias que pertencem a cada cluster de CDP e Lj. Para se poder analisar a dispersão dos clusters no país, e partindo da associação freguesia – concelho, foi possível elaborar um mapa de Portugal continental e ilhas onde foram coloridos os concelhos de acordo com os clusters a que pertencem as suas freguesias. Nos mapas de Portugal seguintes é possível analisar quais os concelhos que têm freguesias de clusters diferentes, ou se no concelho todas as freguesias pertencem ao mesmo cluster, ex.; nos clusters de CDP o concelho de Lisboa tem freguesias que pertencem ao cluster 2 e 4, concelho de Sintra tem freguesias pertencentes aos clusters 2, 3 e 4, inversamente temos o concelho de Sesimbra e Loures onde todas as freguesias pertencem ao mesmo cluster, 4 e 2 respetivamente.

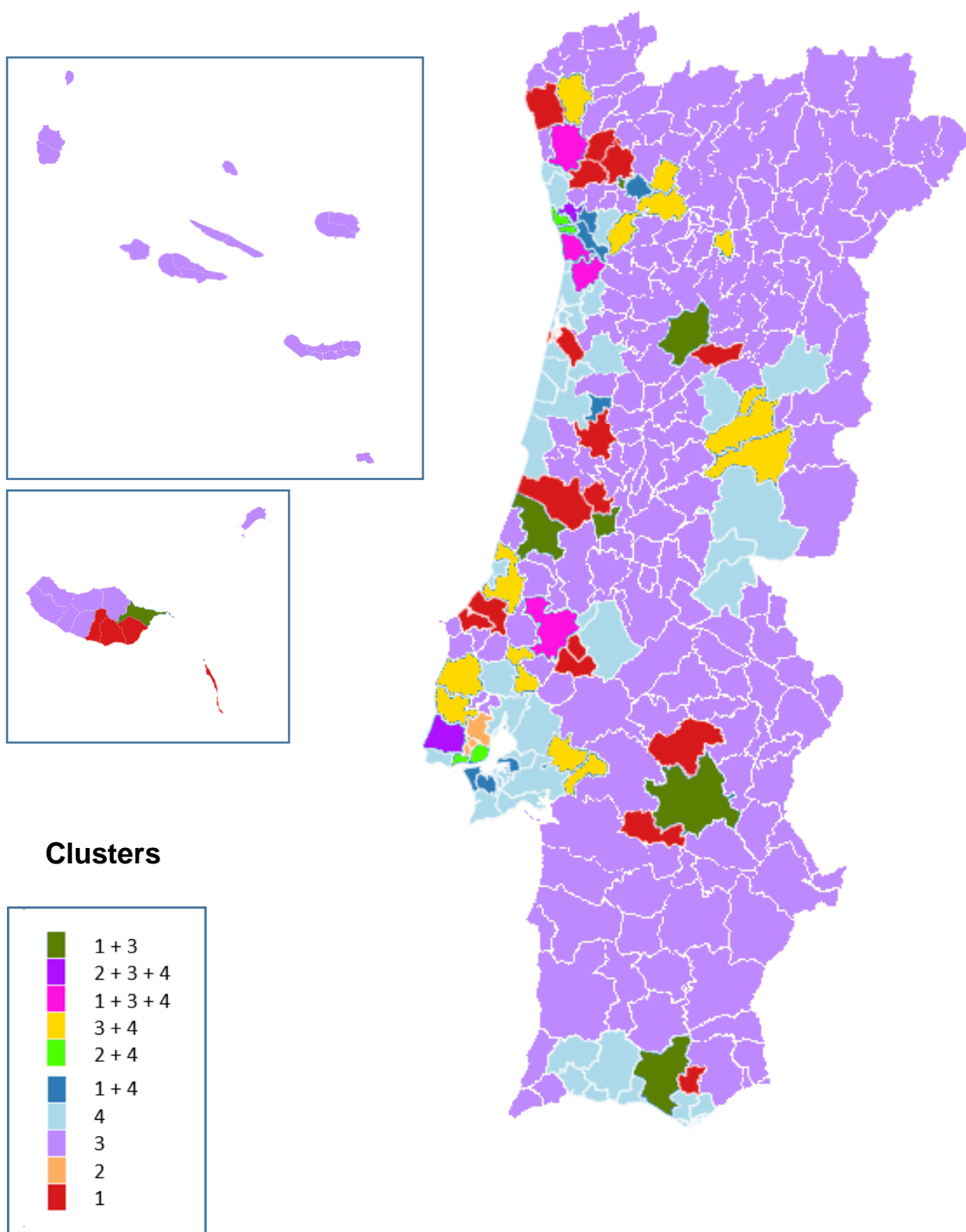


FIGURA 14 - MAPA DE PORTUGAL POR CONCELHOS, CLUSTERS CDP

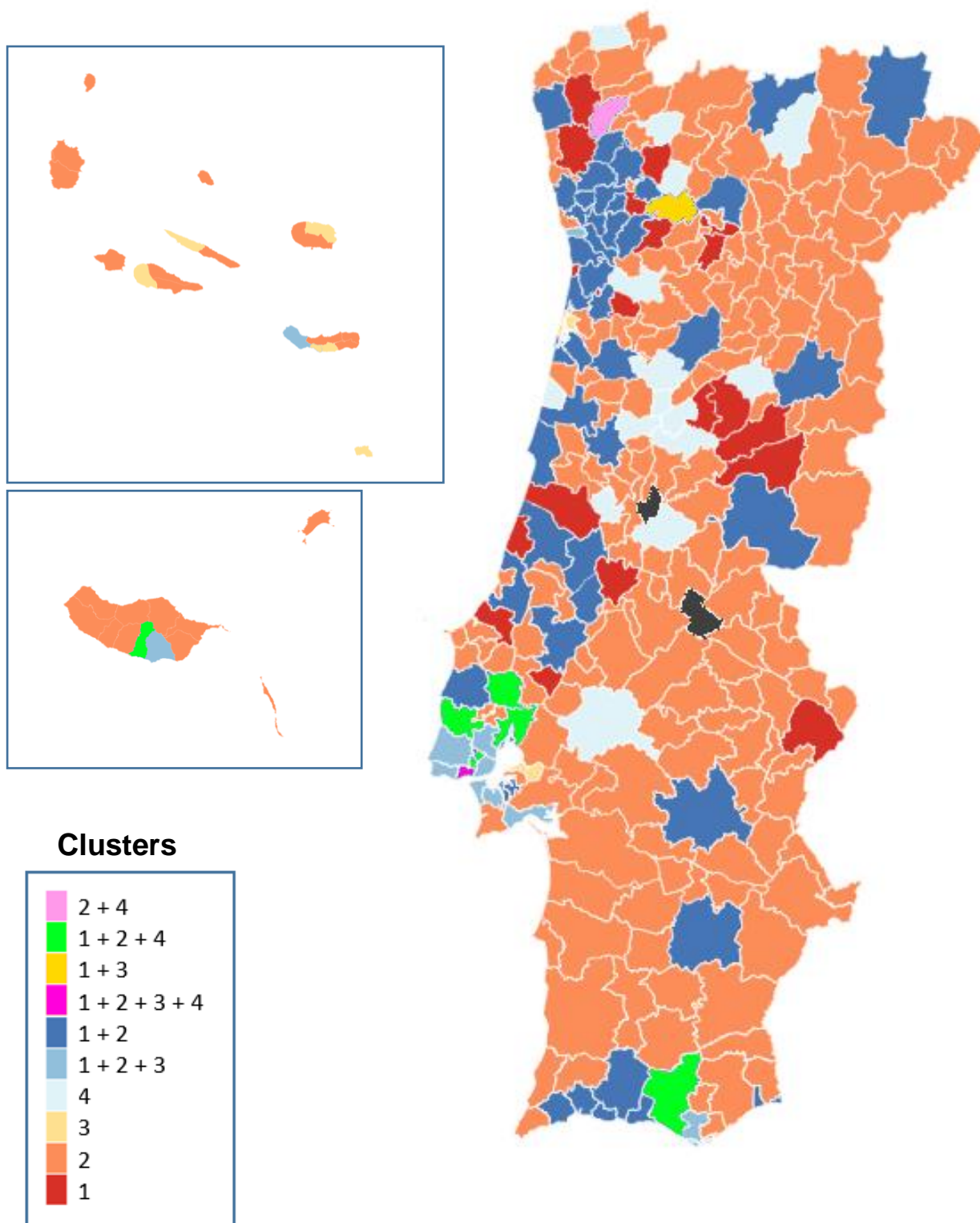


FIGURA 15 - MAPA DE PORTUGAL POR CONCELHOS, CLUSTERS Lj

Neste mapa de Portugal constatamos que existem dois concelhos que não têm Lj, são o concelho do Gavião e Pedrogão Grande, e estão pintados a preto.

4.3. Análise sociodemográfica

Estando os dois primeiros objetivos concluídos, e partindo da lista de freguesias apuradas em resultado das segmentações, foi possível efetuar a análise sociodemográfica com os dados do INE. O resultado foi um quadro com a análise estatística utilizando as variáveis escolhidas, evidenciando as características da população que reside nas freguesias de cada cluster.

4.3.1. Por cluster de CDP

	H	M	0-14	15-24	25-64	65 / +	Alojamentos	nenhuma escolaridade	1_ciclo	2_ciclo
Mean	2 643	2 881	846	622	3 135	921	2 737	432	1 602	604
Median	1 404	1 518	428	328	1 566	536	1 424	265	993	360
Mode	643	2 007	443	148	2 622	293	1 944	352	737	183
Standard Deviation	3 546	3 928	1 177	834	4 334	1 195	3 829	513	1 880	717
Sample Variance	12 576 518	15 429 134	1 385 642	696 117	18 780 868	1 427 892	14 660 840	263 580	3 534 020	514 458
Range	24 287	27 927	7 824	5 629	30 330	9 159	25 963	3 749	14 194	5 082
Minimum	99	109	19	11	87	77	138	24	99	22
Maximum	24 386	28 036	7 843	5 640	30 417	9 236	26 101	3 773	14 293	5 104
Sum	1 020 183	1 112 134	326 523	240 240	1 209 951	355 603	1 056 356	166 735	618 471	233 210
Count	386	386	386	386	386	386	386	386	386	386

	3_ciclo	Ensino secundario	pos-secundario	Superior	desempregados	Total empregada	primario	secundario	terceario	Pensionistas /reformados
Mean	872	924	49	900	367	2 363	38	707	1 618	1 111
Median	443	424	21	275	167	1 149	22	489	650	666
Mode	172	196	11	116	103	637	13	1 384	507	441
Standard Deviation	1 211	1 400	77	1 813	560	3 281	54	733	2 659	1 427
Sample Variance	1 467 700	1 961 017	5 976	3 286 576	313 456	10 764 838	2 891	537 881	7 068 308	2 035 211
Range	8 896	10 140	522	18 724	4 477	22 132	401	4 633	18 224	10 849
Minimum	22	14	0	8	9	52	0	19	29	88
Maximum	8 918	10 154	522	18 732	4 486	22 184	401	4 652	18 253	10 937
Sum	336 591	356 718	18 837	347 273	141 846	912 266	14 806	272 814	624 646	428 795
Count	386	386	386	386	386	386	386	386	386	386

FIGURA 16 - DADOS SOCIODEMOGRAFICOS DO CLUSTER 1 CDP

Neste cluster temos 386 freguesias onde a média de habitantes por freguesia é de 2600 habitantes masculinos e 2800 femininos e com um total de 2 132 317 habitantes, mas com uma grande amplitude de valores desde os 99 até 24000. Habitantes maioritariamente na faixa etária dos 25-64, com escolaridade do 1º ciclo e trabalhadores do setor terciário. Temos também um total de 1 056 356 alojamentos, com uma média 2700 alojamentos. Freguesias com uma média de população elevada, mas claramente com áreas mistas quanto à sua ruralidade.

	H	M	0-14	15-24	25-64	65 / +	Alojamentos	nenhuma escolaridade	1_ciclo	2_ciclo
Mean	12 120	13 844	3 687	2 712	14 432	5 133	13 786	1 765	6 431	2 118
Median	10 674	12 423	3 106	2 401	12 774	4 601	13 109	1 385	5 351	1 703
Mode	N/A	N/A	2 478	N/A	N/A	7 417	N/A	2 829	N/A	N/A
Standard Deviation	7 102	8 032	2 374	1 683	8 621	3 003	7 749	1 113	4 023	1 369
Sample Variance	50 438 911	64 515 886	5 635 924	2 832 515	74 329 404	9 016 904	60 049 792	1 239 160	16 183 569	1 874 270
Range	29 753	33 618	9 336	6 956	34 652	12 449	31 934	5 121	18 302	5 678
Minimum	1 403	1 398	408	293	1 537	543	1 429	226	926	302
Maximum	31 156	35 016	9 744	7 249	36 189	12 992	33 363	5 347	19 228	5 980
Sum	739 326	844 472	224 897	165 442	880 333	313 139	840 976	107 666	392 320	129 212
Count	61	61	61	61	61	61	61	61	61	61

	3_ciclo	Ensino secundario	pos-secundario	Superior	desempregados	Total empregada	primario	secundario	terceario	Pensionistas/reformados
Mean	3 967	4 621	242	6 187	1 717	11 074	34	1 686	9 353	5 895
Median	3 529	3 732	201	4 970	1 354	9 362	29	1 249	8 128	5 201
Mode	N/A	3 257	188	N/A	N/A	N/A	21	N/A	N/A	3 988
Standard Deviation	2 572	2 990	155	4 568	1 138	6 655	30	1 188	5 624	3 454
Sample Variance	6 615 700	8 941 504	23 975	20 868 088	1 295 567	44 284 135	917	1 411 499	31 623 921	11 929 916
Range	10 848	12 285	640	23 410	4 849	26 965	198	4 220	23 293	14 761
Minimum	457	473	25	305	142	1 267	3	278	969	593
Maximum	11 305	12 758	665	23 715	4 991	28 232	201	4 498	24 262	15 354
Sum	242 008	281 864	14 738	377 383	104 715	675 484	2 081	102 849	570 547	359 578
Count	61	61	61	61	61	61	61	61	61	61

FIGURA 17 - DADOS SOCIODEMOGRÁFICOS DO CLUSTER 2 CDP

Este cluster, só com 61 freguesias com um total de 1 583 798, com 840 976 alojamentos e apresenta a média de habitantes mais elevada. Representa as grandes cidades devido à concentração. Caracteriza-se por ter habitantes maioritariamente na faixa etária dos 25-64 mas também com uma média alta de +65 anos, com uma média elevada de habitantes tanto com o 1º ciclo como com o ensino superior, trabalhadores no setor terciário e com uma elevada média de Pensionistas/Reformados.

	H	M	0-14	15-24	25-64	65 / +	Alojamentos	nenhuma escolaridade	1_ciclo	2_ciclo
Mean	776	835	229	175	856	351	999	165	558	183
Median	384	412	96	82	398	215	568	96	334	91
Mode	106	117	17	26	129	116	422	51	165	24
Standard Deviation	1 207	1 309	400	281	1 414	453	1 542	218	712	269
Sample Variance	1 457 004	1 713 359	159 926	78 926	1 998 457	204 774	2 376 339	47 666	506 538	72 449
Range	15 125	16 245	4 759	2 995	18 021	5 600	31 460	2 494	8 699	2 836
Minimum	22	21	2	2	24	10	36	2	25	3
Maximum	15 147	16 266	4 761	2 997	18 045	5 610	31 496	2 496	8 724	2 839
Sum	1 645 767	1 771 457	485 761	371 188	1 815 744	744 529	2 118 116	350 738	1 183 799	388 000
Count	2 121	2 121	2 121	2 121	2 121	2 121	2 121	2 121	2 121	2 121

	3_ciclo	Ensino secundario	pos-secundario	Superior	desempregados	Total empregada	primario	secundario	terceario	Pensionistas /reformados
Mean	244	240	13	170	88	620	40	189	391	401
Median	108	92	5	50	37	257	24	76	139	243
Mode	29	16	1	11	12	55	5	14	39	184
Standard Deviation	416	468	26	415	166	1 085	58	345	774	532
Sample Variance	172 885	219 122	669	171 865	27 624	1 176 736	3 401	119 049	599 174	283 521
Range	5 536	6 257	348	5 970	2 047	13 571	1 191	6 682	8 965	7 480
Minimum	4	2	0	0	0	14	0	1	8	14
Maximum	5 540	6 259	348	5 970	2 047	13 585	1 191	6 683	8 973	7 494
Sum	518 326	509 331	26 644	359 529	187 621	1 315 826	85 853	400 061	829 910	851 367
Count	2 121	2 121	2 121	2 121	2 121	2 121	2 121	2 121	2 121	2 121

FIGURA 18 - DADOS SOCIODEMOGRÁFICOS DO CLUSTER 3 CDP

Sendo este o cluster maior, com 2 121 freguesias, um total de 3 417 224 habitantes e 2 118 116 alojamentos, tem médias e modas muito baixas tanto de população como de alojamentos por freguesia. São as freguesias com maiores áreas rurais. Este cluster caracteriza-se por ter uma população na faixa etária dos 25-64 anos, uma escolaridade maioritariamente do 1º ciclo, trabalhadores do setor terciário, mas com valor elevado de pensionistas/reformados, maior que os trabalhadores do sector terciário e quase equiparado ao total de empregados.

	H	M	0-14	15-24	25-64	65 / +	Alojamentos	nenhuma escolaridade	1_ciclo	2_ciclo
Mean	3 438	3 783	1 106	771	4 042	1 302	3 922	549	1 958	704
Median	1 121	1 193	373	274	1 284	440	1 099	206	829	322
Mode	978	845	208	17	216	345	742	172	377	86
Standard Deviation	5 552	6 249	1 843	1 247	6 676	2 159	6 442	821	2 756	999
Sample Variance	30 827 880	39 044 284	3 397 627	1 554 608	44 573 079	4 662 662	41 496 946	674 357	7 593 431	998 819
Range	31 530	34 573	12 375	7 543	38 118	13 973	38 657	5 131	14 236	6 085
Minimum	70	77	6	9	61	42	104	14	70	8
Maximum	31 600	34 650	12 381	7 552	38 179	14 015	38 761	5 145	14 306	6 093
Sum	1 928 900	2 122 024	620 265	432 596	2 267 547	730 519	2 200 249	308 010	1 098 524	395 102
Count	561	561	561	561	561	561	561	561	561	561

	3_ciclo	Ensino secundario	pos-secundario	Superior	desempregados	Total empregada	primario	secundario	terceario	Pensionistas /reformados
Mean	1 158	1 293	67	1 306	474	3 068	58	746	2 265	1 510
Median	379	323	17	201	127	987	26	400	490	513
Mode	233	166	2	26	79	311	18	17	192	259
Standard Deviation	1 917	2 346	123	2 972	834	5 113	103	949	4 271	2 471
Sample Variance	3 676 246	5 505 176	15 045	8 833 411	695 857	26 141 635	10 702	900 362	18 245 546	6 106 605
Range	12 930	15 933	791	23 714	4 941	30 314	1 535	6 349	23 919	15 649
Minimum	11	4	0	1	3	40	0	8	12	50
Maximum	12 941	15 937	791	23 715	4 944	30 354	1 535	6 357	23 931	15 699
Sum	649 717	725 314	37 570	732 924	265 819	1 721 236	32 486	418 275	1 270 471	846 836
Count	561	561	561	561	561	561	561	561	561	561

FIGURA 19 - DADOS SOCIODEMOGRÁFICOS DO CLUSTER 4 CDP

Este cluster, não sendo o maior em número de freguesias, 561, é o maior em população total 4 050 924 e em alojamentos 2 200 249, Este cluster representa grandes concentrações populacionais, onde se destaca algum nivelamento ao nível da escolaridade, uma forte representação do setor terciário e, mais uma vez, uma média elevada de Pensionistas/reformados.

A análise sociodemográfica aos clusters de CDP permite concluir que existe uma grande discrepância quanto ao nível médio de habitantes, sendo possível também distinguir os cluster mais rurais dos urbanos. A faixa etária predominante é 25-64, o nível de escolaridade alterna entre o 1º ciclo e o ensino superior, sendo que, nas cidades o nível de escolaridade é maioritariamente o ensino superior, contando também com uma grande faixa de Pensionistas/reformados.

4.3.2. Por cluster de Lj

	<i>H</i>	<i>M</i>	<i>0-14</i>	<i>15-24</i>	<i>25-64</i>	<i>65 / +</i>	<i>Alojamentos</i>	<i>nenhuma escolaridade</i>	<i>1_ciclo</i>	<i>2_ciclo</i>
Mean	1 537	1 698	482	357	1 806	591	1 781	248	900	330
Median	499	531	146	117	557	198	531	98	366	131
Mode	426	50	12	9	53	138	156	107	120	22
Standard Deviation	3 558	4 021	1 149	819	4 272	1 416	4 332	523	1 797	657
Sample Variance	12 661 925	16 167 338	1 320 660	670 851	18 252 795	2 003 809	18 761 991	273 020	3 230 814	431 515
Range	31 565	34 612	12 380	7 548	38 149	13 996	38 709	4 977	14 260	6 089
Minimum	35	39	1	4	31	19	52	6	33	4
Maximum	31 600	34 650	12 381	7 552	38 179	14 015	38 761	4 982	14 293	6 093
Sum	1 862 894	2 058 570	583 778	433 124	2 188 749	715 816	2 158 583	300 467	1 091 101	400 165
Count	1 212	1 212	1 212	1 212	1 212	1 212	1 212	1 212	1 212	1 212

	<i>3_ciclo</i>	<i>Ensino secundario</i>	<i>pos-secundario</i>	<i>Superior</i>	<i>desempregados</i>	<i>Total empregada</i>	<i>primario</i>	<i>secundario</i>	<i>terceario</i>	<i>Pensionistas/reformados</i>
Mean	515	557	29	576	209	1 364	24	370	970	692
Median	160	136	7	82	56	409	13	151	203	231
Mode	15	69	1	4	5	34	6	9	135	134
Standard Deviation	1 238	1 490	78	1 829	527	3 274	55	688	2 672	1 618
Sample Variance	1 532 892	2 218 768	6 080	3 346 884	277 720	10 717 752	3 077	473 848	7 140 573	2 617 119
Range	12 937	15 934	791	20 206	4 943	30 339	1 535	6 683	23 926	15 681
Minimum	5	4	0	1	1	15	0	1	6	19
Maximum	12 941	15 937	791	20 207	4 944	30 354	1 535	6 683	23 931	15 699
Sum	624 655	675 429	34 913	697 852	252 900	1 653 094	29 068	448 814	1 175 210	838 799
Count	1 212	1 212	1 212	1 212	1 212	1 212	1 212	1 212	1 212	1 212

FIGURA 20 - DADOS SOCIODEMOGRÁFICOS DO CLUSTER 1 Lj

Neste primeiro cluster temos valores médios elevados na população bem como na habitação. A faixa etária predominante é dos 25-64 e o nível de escolaridade com mais indivíduos é o 1º ciclo, sendo que o 3º ciclo, ensino secundário e superior têm valores muito semelhantes.

	<i>H</i>	<i>M</i>	<i>0-14</i>	<i>15-24</i>	<i>25-64</i>	<i>65 / +</i>	<i>Alojamentos</i>	<i>nenhuma escolaridade</i>	<i>1_ciclo</i>	<i>2_ciclo</i>
Mean	1 291	1 410	401	289	1 491	520	1 525	232	797	272
Median	370	400	98	80	393	192	496	88	305	90
Mode	106	138	14	16	102	65	218	33	127	12
Standard Deviation	3 187	3 556	1 057	724	3 839	1 179	3 624	488	1 657	597
Sample Variance	10 159 844	12 641 604	1 117 293	524 167	14 736 766	1 390 469	13 130 350	238 355	2 746 271	356 253
Range	31 578	34 995	12 380	7 550	38 155	12 982	38 725	5 345	19 203	6 090
Minimum	22	21	1	2	24	10	36	2	25	3
Maximum	31 600	35 016	12 381	7 552	38 179	12 992	38 761	5 347	19 228	6 093
Sum	3 316 146	3 622 615	1 030 840	741 347	3 831 231	1 335 352	3 918 252	595 607	2 048 222	699 780
Count	2 569	2 569	2 569	2 569	2 569	2 569	2 569	2 569	2 569	2 569

	<i>3_ciclo</i>	<i>Ensino secundario</i>	<i>pos-secundario</i>	<i>Superior</i>	<i>desempregados</i>	<i>Total empregada</i>	<i>primario</i>	<i>secundario</i>	<i>terceario</i>	<i>Pensionistas/reformados</i>
Mean	423	460	24	424	170	1 117	37	273	807	602
Median	111	95	5	53	38	269	18	86	149	222
Mode	39	16	1	9	5	55	5	11	35	76
Standard Deviation	1 104	1 325	70	1 570	477	2 947	69	561	2 424	1 373
Sample Variance	1 218 070	1 754 451	4 924	2 463 723	227 056	8 685 828	4 785	314 904	5 874 864	1 885 008
Range	12 937	15 935	791	23 715	4 991	30 340	1 535	6 357	24 257	15 340
Minimum	4	2	0	0	0	14	0	1	6	14
Maximum	12 941	15 937	791	23 715	4 991	30 354	1 535	6 357	24 262	15 354
Sum	1 087 591	1 182 986	61 971	1 090 301	435 894	2 868 644	94 616	700 717	2 073 299	1 546 943
Count	2 569	2 569	2 569	2 569	2 569	2 569	2 569	2 569	2 569	2 569

FIGURA 21 - DADOS SOCIODEMOGRÁFICOS DO CLUSTER 2 Lj

Este cluster é o maior ao nível de freguesias e de população, mas ao nível de médias é muito semelhante ao anterior. Em valores absolutos existem diferenças mas devido ao número de freguesias, quase o dobro, as diferenças tendem a desaparecer. Aqui situam-se as freguesias maioritariamente rurais.

	<i>H</i>	<i>M</i>	<i>0-14</i>	<i>15-24</i>	<i>25-64</i>	<i>65 / +</i>	<i>Alojamentos</i>	<i>nenhuma escolaridade</i>	<i>1_ciclo</i>	<i>2_ciclo</i>
Mean	2 327	2 600	750	538	2 755	885	2 574	349	1 255	441
Median	424	448	146	106	475	138	459	88	303	111
Mode	410	135	29	74	692	47	94	25	195	24
Standard Deviation	5 066	5 786	1 591	1 122	6 139	2 133	5 792	728	2 521	853
Sample Variance	25 663 352	33 475 416	2 531 421	1 259 089	37 684 521	4 547 720	33 545 882	530 280	6 354 135	727 735
Range	28 083	31 367	8 975	5 965	34 911	13 996	29 355	4 595	13 820	4 652
Minimum	49	62	10	11	60	19	67	6	46	10
Maximum	28 131	31 428	8 984	5 976	34 970	14 015	29 422	4 600	13 865	4 662
Sum	416 515	465 428	134 179	96 306	493 135	158 327	460 748	62 430	224 667	78 991
Count	179	179	179	179	179	179	179	179	179	179

	<i>3_ciclo</i>	<i>Ensino secundario</i>	<i>pos-secundario</i>	<i>Superior</i>	<i>desempregados</i>	<i>Total empregada</i>	<i>primario</i>	<i>secundario</i>	<i>terceario</i>	<i>Pensionistas/reformados</i>
Mean	770	896	47	1 041	313	2 094	38	361	1 696	1 016
Median	133	123	5	67	50	357	20	88	213	158
Mode	30	32	1	15	22	55	18	273	47	57
Standard Deviation	1 713	2 135	110	2 760	718	4 721	55	704	4 023	2 446
Sample Variance	2 935 229	4 559 022	12 159	7 618 922	514 887	22 289 816	3 028	494 931	16 185 187	5 984 254
Range	9 400	12 109	658	16 102	3 641	28 189	451	3 924	24 245	15 682
Minimum	12	14	0	2	1	44	2	9	18	18
Maximum	9 412	12 123	658	16 104	3 641	28 232	453	3 932	24 262	15 699
Sum	137 771	160 326	8 340	186 302	56 015	374 896	6 738	64 619	303 537	181 779
Count	179	179	179	179	179	179	179	179	179	179

FIGURA 22 - DADOS SOCIODEMOGRÁFICOS DO CLUSTER 3 LJ

A média de população neste cluster é a mais elevada, bem como a de alojamentos, em contraponto é a menor em número de freguesias. A faixa etária com maiores valores é 25-64 e o nível de escolaridade onde a média é mais elevada é o 1º ciclo. Freguesias com grande densidade populacional, média elevada de trabalhadores do sector terciário e Pensionistas/reformados.

	<i>H</i>	<i>M</i>	<i>0-14</i>	<i>15-24</i>	<i>25-64</i>	<i>65 / +</i>	<i>Alojamentos</i>	<i>nenhuma escolaridade</i>	<i>1_ciclo</i>	<i>2_ciclo</i>
Mean	498	548	144	109	551	244	723	110	366	109
Median	248	279	61	51	265	140	342	57	219	58
Mode	103	131	60	31	199	84	343	45	328	53
Standard Deviation	1 245	1 381	437	272	1 484	455	2 111	231	619	212
Sample Variance	1 550 806	1 906 918	190 551	73 832	2 201 468	207 070	4 457 548	53 155	383 001	44 960
Range	17 012	19 198	6 400	3 593	20 648	5 592	31 422	2 715	6 844	2 401
Minimum	40	40	6	6	29	17	75	4	26	3
Maximum	17 051	19 237	6 405	3 598	20 676	5 609	31 496	2 719	6 870	2 404
Sum	169 440	186 490	48 937	36 904	187 188	82 902	245 671	37 567	124 484	37 133
Count	340	340	340	340	340	340	340	340	340	340

	<i>3_ciclo</i>	<i>Ensino secundario</i>	<i>pos-secundario</i>	<i>Superior</i>	<i>desempregados</i>	<i>Total empregada</i>	<i>primario</i>	<i>secundario</i>	<i>terceario</i>	<i>Pensionistas /reformados</i>
Mean	150	157	9	120	54	396	23	114	259	273
Median	70	59	3	35	23	170	12	61	94	158
Mode	41	21	2	12	4	131	8	28	21	65
Standard Deviation	396	510	29	637	175	1 166	42	192	975	520
Sample Variance	157 038	259 601	853	405 426	30 596	1 358 565	1 749	36 909	950 394	270 874
Range	4 941	6 618	385	10 971	2 065	16 648	489	2 061	14 556	6 766
Minimum	8	7	0	3	2	19	0	3	9	18
Maximum	4 948	6 624	385	10 973	2 067	16 667	489	2 064	14 565	6 784
Sum	51 057	53 337	2 991	40 955	18 487	134 643	7 960	38 604	88 079	92 868
Count	340	340	340	340	340	340	340	340	340	340

FIGURA 23 - DADOS SOCIODEMOGRÁFICOS DO CLUSTER 4 LJ

Este cluster apresenta valores baixos nas médias de população, bem como no valor total de habitantes. A faixa etária com maiores valores é 25-64 e o nível de escolaridade onde a média é mais elevada é o 1º ciclo.

5. Conclusões

Neste trabalho foi efetuada uma segmentação de CDP e LJ de acordo com as suas características, e identificadas as respetivas freguesias, com o objetivo de produzir uma caracterização sociodemográfica da população cliente dos CTT.

Na segmentação dos CDP foi possível criar clusters que se diferenciaram por tipo de correio e ruralidade das freguesias. Em alguns clusters foi também possível identificar padrões de serviço associados a empresas.

Quanto às LJ, os segmentos resultantes não foram capazes de espelhar com tanta evidência quais os que pertencem a meios rurais ou urbanos, foi possível sim reconhecer padrões de tipo de serviço, uns direcionados para empresas e outros para particulares. Ainda assim ficaram agregadas as LJ que, devido à média de clientes/dia e alguns tipos de serviços, representam as cidades, por contraponto com as LJ que têm como clientes muitas empresas.

Dos clusters obtidos foi possível uma representação gráfica por concelhos no mapa de Portugal, observando-se genericamente uma semelhança na dispersão dos clusters de CDP e Lj, por existir uma maior concentração de população no litoral norte, em comparação com o resto do país. Contudo, verifica-se que no norte e centro do país, ao nível das Lj existe uma alternância de clusters, o que significa que as características das Lj são diferentes dos CDP.

Na análise sociodemográfica efetuada sobre a população das freguesias de cada cluster, obteve-se nos CDP uma separação mais nítida entre meios rurais, com baixa densidade populacional, e meios urbanos. Já nas Lj essa separação não é tão clara, pelo facto de os valores médios serem semelhantes. A faixa etária predominante em todos os clusters – CDP e Lj -, foi entre os 25-64 anos, o nível de escolaridade situou-se entre o 1º ciclo e o ensino superior, sendo que o sector que emprega mais população é o setor terciário. Nos clusters onde se identifica grande densidade populacional, também se observam valores elevados de Pensionistas/reformados.

Da análise global aos resultados obtidos conclui-se ainda que os dados utilizados para a segmentação dos CDP conseguiram ser mais diferenciadores que os das Lj, o que se deve ao facto de as variáveis associadas aos CDP serem mais discriminativas, ao contrário das relativas às Lj, que eram demasiado agregadas.

6. Recomendações para Trabalhos Futuros

Com o aproximar dos próximos censos em 2021 e com a entrada em serviço do Banco CTT, S.A., seria interessante verificar se ocorrem variações na forma como os clusters, Lj ou CDP, estão dispersos no mapa de Portugal.

A utilização de outras variáveis, menos agregadoras, de forma a avaliar o seu impacto no resultado final.

Este trabalho pode servir também como base de comparação com futuras análises de Lj e CDP, após a entrada de um novo produto ou serviço no negócio dos CTT.

7. Bibliografia

- ANACOM. (2016). *Serviços Postais - 3.º trimestre de 2016*. Retrieved from <http://www.anacom.pt/render.jsp?contentId=1401255>
- ANACOM. (2014). ANACOM - Serviços Postais - 4.º trimestre de 2014. Retrieved May 7, 2015, from <http://www.anacom.pt/render.jsp?contentId=1348279#.VUts466eCg>
- Baçaõ, F., Lobo, V., & Painho, M. (2005). The self-organizing map, the Geo-SOM, and relevant variants for geosciences. *Computers and Geosciences*, 31(2), 155–163. <https://doi.org/10.1016/j.cageo.2004.06.013>
- Benítez, I., Quijano, A., Díez, J., & Delgado, I. (2014). Dynamic clustering segmentation applied to load profiles of energy consumption from Spanish customers. *International Journal of Electrical Power and Energy Systems*, 55, 437–448. <https://doi.org/10.1016/j.ijepes.2013.09.022>
- Chicco, G., Napoli, R., & Piglion, F. (2006). Comparisons among clustering techniques for electricity customer classification. *IEEE Transactions on Power Systems*, 21(2), 933–940. <https://doi.org/10.1109/TPWRS.2006.873122>
- Henriques, R. (2014). A observação da evolução nos padrões espaciais e sociodemográficos através dos self-organizing maps : o caso da Área Metropolitana de Lisboa – Norte.
- Henriques, R. A. P. (2011, June 6). Artificial Intelligence in geospatial analysis: applications of self-organizing maps in the context of geographic information science. Retrieved from <http://run.unl.pt/handle/10362/5723>
- Henriques, R. A. P. (2006, February 15). CARTO-SOM - Cartogram creation using self-organizing maps. Retrieved from <http://run.unl.pt/handle/10362/3641>
- Henriques, R., Bacao, F., & Lobo, V. (2012). Exploratory geospatial data analysis using the GeoSOM suite. *Computers, Environment and Urban Systems*, 36(3), 218–232. <https://doi.org/10.1016/j.compenvurbsys.2011.11.003>
- Henriques, R., Lobo, V., & Bacao, F. (2010). Artificial Intelligence in Geospatial Analysis : applications of Self-Organizing Maps in the context of Geographic Information Science, (March).
- Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9). <https://doi.org/10.1109/5.58325>
- Kohonen, T. (1997). Exploration of Very Large Databases by Self-Organizing Maps. *Proceedings of ICNN'97, International Conference on Neural Networks*, PL1--PL6.
- Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks : The Official Journal of the International Neural Network Society*, 37, 52–65. <https://doi.org/10.1016/j.neunet.2012.09.018>

- Kohonen, T. (2001). *Self-Organizing Maps. Springer Series in Information Sciences* (Vol. 30). <https://doi.org/10.1007/978-3-642-56927-2>
- Kohonen, T., Oja, M., Kaski, S., & Somervuo, P. (1990). Self-organizing map. *Proceedings of the IEEE*, 111–116. Retrieved from https://scholar.google.fr/citations?view_op=view_citation&continue=/scholar%3Fhl%3Dfr%26as_sdt%3D0,5%26scilib%3D1&citilm=1&citation_for_view=YbDdHsMAAAAJ:YsMSGlbeyi4C&hl=fr&oi=p
- Li, Y., & Lin, F. (2008). Customer segmentation analysis based on SOM clustering. *Proceedings of 2008 IEEE International Conference on Service Operations and Logistics, and Informatics, IEEE/SOLI 2008, 1*, 15–19. <https://doi.org/10.1109/SOLI.2008.4686353>
- López, J. J., Aguado, J. A., Martín, F., Munoz, F., Rodríguez, A., & Ruiz, J. E. (2011). Hopfield-K-Means clustering algorithm: A proposal for the segmentation of electricity customers. *Electric Power Systems Research*, 81(2), 716–724. <https://doi.org/10.1016/j.epsr.2010.10.036>
- Sánchez, I. B., Espinós, I. D., Sarrión, L. M., López, A. Q., & Burgos, I. N. (2009). Clients segmentation according to their domestic energy consumption by the use of self-organizing maps. *2009 6th International Conference on the European Energy Market, EEM 2009*, 1–6. <https://doi.org/10.1109/EEM.2009.5207172>
- Universal Postal Union. (2015). *Universal Postal Union – 2015 Results*. Retrieved from <http://www.upu.int/en/resources/postal-statistics/2015-results.html>
- Universal Postal Union. (2013.). *Universal Postal Union – 2013 Results*. Retrieved May 7, 2015, from <http://www.upu.int/en/resources/postal-statistics/2013-results.html>
- Verdú, S. V., García, M. O., Senabre, C., Marín, A. G., & Franco, F. J. G. (2006). Classification, filtering, and identification of electrical customer load patterns through the use of self-organizing maps. *IEEE Transactions on Power Systems*, 21(4), 1672–1682. <https://doi.org/10.1109/TPWRS.2006.881133>
- Vesanto, J. (1999). SOM-based data visualization methods. *Intelligent Data Analysis*, 3(2), 111–126. <https://doi.org/10.3233/IDA-1999-3203>
- Vesanto, J., Alhoniemi, E., & Member, S. (2000). Clustering of the Self-Organizing Map, *11*(3), 586–600. <https://doi.org/10.1109/72.846731>
- Vesanto, J., Himberg, J., Alhoniemi, E., & Parhankangas, J. (2000). SOM Toolbox for Matlab 5. *Technical Report A57*, 2(0), 59. <https://doi.org/http://www.cis.hut.fi/somtoolbox/package/papers/techrep.pdf>

8. Anexos

8.1. Anexo 1

Como não é possível visualizar na matriz U todos os códigos de cada posição, em baixo está a lista com todos os códigos de cada posição e os que são visíveis identificados a amarelo, 1º nível.

1º Nível	2º Nível	3º Nível	1º Nível	2º Nível	3º Nível	1º Nível	2º Nível	3º Nível
222	6300		3460	3360		5140		
999			3500			5200		
1050			3550	4860		5210	4850	
1300			3560			5300		
1500			3600			5340	5320	
1600	1200		3620	5120		5350	4960	
1800			3670			5370		
2000			3700			5400		
2040	6230		3720	3060	3080	5430	3640	
2070	3740		3730			5450		
2125			3750			5470		
2140			3770			6000	2765	8500
2300			3780			6060		
2350			3800			6090	6150	
2400			3830			6100	6290	
2425			3850	2495		6120	7220	
2430	8900		3880			6160		
2460			4000			6200		
2475	7450		4100			6250		
2500			4200			6320	2550	6400
2520	2330		4400			6355		
2540			4415			6420		
2565			4420			6440	5230	
2580			4430			7000		
2600	8125	3520	4435			7050		
2615			4445			7080		
2625	2855		4450	1700		7170	7160	
2630			4470			7200	3440	
2665			4480			7250		
2670			4490			7300		
2685	2675		4500	3860		7350		
2700	1100		4520			7400		
2710			4540	4690		7440		
2715			4550			7460		
2735			4560	2640		7490	7470	
2745			4590	2100		7520		
2780	2750		4600			7540	7500	
2785			4610			7570		
2795			4620	2530	4575	7580		
2800			4630			7600		
2825	2810		4640			7630	7830	
2830			4660			7700	7900	
2835			4700	3000		7780	7750	
2840	2775		4730			7800		
2845			4740			7860		

2860			4750	4580		7940		
2870			4780	2480		8000		
2900	2910		4785			8100		
2950			4800			8200		
2970			4820			8400		
3050	6270		4830			8600		
3100			4890	4950		8700		
3150			4900	4760		8800	7100	7960
3200			4910	9800	9600	9000		
3220			4980	4930		9200		
3300			4990	2560		9360		
3330			5000	2200		9400		
3400			5050	3140		9500		
3420			5070	5160		9650		
			5100	2490		9700	3260	
						9760		
						9880	9580	
						9900		
						9950		
						9970		
						9980	6430	

8.2. Anexo 2

Tabela de conversão de ID em código e designação da Lj.

id	código	Designação		id	código	Designação		id	código	Designação
1	8810747	RESTAURADO RES (LISBOA)		202	8810044	POUSADA DE SARAMAGOS		404	8810927	ALJUSTREL
2	8810782	CABO RUIVO (LISBOA)		203	8810151	PÓVOA DE VARZIM		405	8810703	SEIXAL
3	8810757	CHIADO (LISBOA)		204	8810158	AMARANTE		406	8810704	ALDEIA DE PAIO PIRES
4	8810765	AMOREIRAS (LISBOA)		205	8810570	ALVERCA		407	8811032	SANTA CRUZ (MADEIRA)
5	8810950	FARO		206	8810939	TAVIRA		408	8810914	SERPA
6	8811474	CORTE INGLES (LISBOA)		207	8817048	CASCAIS II		409	8810922	CASTRO VERDE
7	8810781	AEROPORTO (LISBOA)		208	8810751	MORAIS SOARES (LISBOA)		410	8810437	TONDELA
8	8810194	BOAVISTA (PORTO)		209	8810830	MEM MARTINS		411	8810909	MOURA
9	8811029	ZARCO (FUNCHAL)		210	8810625	SANTARÉM		412	8810294	SEVER DO VOUGA
10	8810033	ROTUNDA (BRAGA)		211	8810708	BARREIRO		413	8810417	SANTA COMBA DÃO
11	8811328	MARRAZES		212	8810623	ALCANENA		414	8810726	MONTE DA CAPARICA
12	8810034	MAXIMINOS (BRAGA)		213	8811462	SANTA MARTA DO PINHAL		415	8810300	MOURISCA DO VOUGA
13	8810761	RATO (LISBOA)		214	8810806	VENDA NOVA (AMADORA)		416	8810088	MOGADOURO
14	8810793	S. DOMINGOS DE BENFICA (LISBOA)		215	8810200	AMIAL (PORTO)		417	8810060	PÓVOA DE LANHOSO
15	8810983	TEIXEIRA GOMES (PORTIMÃO)		216	8810072	BRAGANÇA		418	8810919	SANTIAGO DO CACÉM
16	8810037	BARCELOS		217	8810268	AVENIDA (AVEIRO)		419	8810677	SALVATERRA DE MAGOS

17	8810772	CINCO DE OUTUBRO (LISBOA)		218	8810601	MAFRA		420	8810628	ALCANEDE
18	8811353	COLOMBO (LISBOA)		219	8810019	VALENÇA		421	8811046	PONTA DO SOL
19	8810746	SOCORRO (LISBOA)		220	8810390	SANTA CRUZ (COIMBRA)		422	8810475	S. PEDRO DO SUL
20	8810195	PEDRO HISPANO (PORTO)		221	8810582	MOSCAVIDE (LISBOA)		423	8810249	VILA DE CUCUJÃES
21	8810036	VILA NOVA DE FAMALICÃO		222	8810139	FREIXEIRO		424	8810251	CESAR
22	8810210	MARQUÊS DE POIMBAL (PORTO)		223	8810686	PALMELA		425	8810488	MOIMENTA DA BEIRA
23	8810131	MAIA		224	8810563	CASTELO BRANCO		426	8810320	NAZARÉ
24	8810810	CASCAIS		225	8810573	PORTELA		427	8811031	PORTO SANTO
25	8811407	MONTÉ BELO (SETÚBAL)		226	8810049	GUIMARÃES		428	8810313	OIÃ
26	8810883	ROSSIO (ÉVORA)		227	8810228	ARCOZELO (VILA NOVA DE GAIA)		429	8810553	PROENÇA-A-NOVA
27	8810641	FÁTIMA		228	8810838	MASSAMÁ		430	8810442	CARREGAL DO SAL
28	8810244	S. JOÃO DA MADEIRA		229	8810754	OLAIAS (LISBOA)		431	8810224	CASTELO DE PAIVA
29	8811409	TECMAIA		230	8817035	LOJA DO CIDADÃO ODIVELAS		432	8810439	NELAS
30	8810823	ALCABIDECHÉ		231	8810834	QUELUZ		433	8810385	SOURE
31	8810769	PICOAS (LISBOA)		232	8810749	ARROIOS (LISBOA)		434	8810515	PINHEL
32	8810032	AVENIDA (BRAGA)		233	8810780	XABREGAS (LISBOA)		435	8810256	ESTARREJA
33	8810231	LOUROSA		234	8810941	OLHÃO		436	8810052	PEVIDÉM
34	8810206	BONFIM (PORTO)		235	8810602	LOURINHÃ		437	8810908	VIDIGUEIRA
35	8811478	ABÓBODA		236	8810399	BAIRRO DA ESTAÇÃO (FIGUEIRA DA FOZ)		438	8810888	CAMPO MAIOR
36	8810958	ALMANCIL		237	8810689	N. S. DA CONCEIÇÃO (SETÚBAL)		439	8810062	VIEIRA DO MINHO
37	8811477	MADEIRA SHOPPING		238	8811454	AGUALVA-CACÉM		440	8810461	GOUVEIA
38	8810979	GIL EANES (PORTIMÃO)		239	8810595	PENICHE		441	8810048	RIBA DE AVE
39	8810286	ÁGUEDA		240	8811491	ABRAVESES (VISEU)		442	8810357	MIRANDA DO CORVO
40	8810325	BENEDITA		241	8810588	SANTO ANTONIO DOS CAVALEIROS		443	8811001	LAGOA (S. MIGUEL)
41	8810173	ERMESINDE		242	8810725	COSTA DA CAPARICA		444	8810468	CINFÃES
42	8810962	CERRO DA ALAGOA (ALBUFEIRA)		243	8810178	VALONGO		445	8810455	TRANCOSO
43	8810980	LAGOS		244	8810853	PORTALEGRE		446	8811053	ESTREITO CÂMARA DE LOBOS
44	8810956	VILAMOURA		245	8810155	LOUSADA		447	8810680	AZAMBUJA
45	8810755	CASAL RIBEIRO (LISBOA)		246	8810001	VIANA DO CASTELO		448	8810402	BUARCOS (VILA)
46	8810207	EÇA DE QUEIROS (PORTO)		247	8810820	PAÇO DE ARCOS		449	8811049	SANTANA
47	8811351	VALE PARAÍSO (ALBUFEIRA)		248	8810948	LOULÉ		450	8810898	ALCACER DO SAL
48	8810599	TORRES VEDRAS		249	8810322	ALCOBAÇA		451	8810422	TÁBUA
49	8811403	GARE DO ORIENTE (LISBOA)		250	8810627	CARTAXO		452	8810092	VINHAI
50	8810740	SANTA JUSTA (LISBOA)		251	8817002	ZONA INDUSTRIAL DE OLHÃO		453	8810923	MÉRTOLA
51	8810211	VILA NOVA DE GAIA		252	8810250	ARRIFANA		454	8811020	HORTA

52	8811042	CALÓUSTE GULBENKIAN (FUNCHAL)		253	8810361	VILA NOVA DE POIARES		455	8810089	ALFANDEGA DA FÉ
53	8811034	MERCADO (FUNCHAL)		254	8810235	FIÃES		456	8810538	IDANHA-A-NOVA
54	8810811	MARGINAL (CASCAIS)		255	8810946	S. BRÁS DE ALPORTEL		457	8810984	ALJEZUR
55	8810144	TROFA		256	8810051	CALDAS DAS TAIPAS		458	8810658	ABRANTES
56	8810768	ENTRECAMPO S (LISBOA)		257	8811400	PAX JULIA (BEJA)		459	8810319	PATAIAS
57	8810774	ALVALADE (LISBOA)		258	8811402	CANIDelo (VILA NOVA DE GAIA)		460	8810989	VILA DO BISPO
58	8810707	PAIVAS (SEIXAL)		259	8811355	PORTO SALVO		461	8810580	SOBRAL DE MONTE AGRÃO
59	8810157	FELGUEIRAS		260	8811414	ALFENA		462	8810645	CAXARIAS-NORTE
60	8810169	GONDOMAR		261	8817014	NUNO ÁLVARES (CHAVES)		463	8810613	ALHANDRA
61	8810794	BENFICA (LISBOA)		262	8810809	CASAL DE S. BRÁS (AMADORA)		464	8810166	VILA MEÃ
62	8810795	LUMIAR (LISBOA)		263	8810267	ÍLHAVO		465	8810933	VILA NOVA DE MILFONTES
63	8810096	VILA REAL		264	8810701	TORRE DA MARINHA		466	8810581	ARRUDA DOS VINHOS
64	8810431	UISEU		265	8810901	BEJA		467	8811017	SANTA CRUZ DA GRACIOSA
65	8810172	RIO TINTO		266	8811419	BAGUIM DO MONTE		468	8810499	VILA NOVA FOZ CÔA
66	8810198	PEREIRO (PORTO)		267	8811043	PENTEADA (FUNCHAL)		469	8810074	CARRAZEDA ANSIÃES
67	8817001	ALGUEIRÃO		268	8810285	VALE DE CAMBRA		470	8810479	CASTRO DAIRE
68	8810177	S. CAETANO (GONDOMAR)		269	8810709	QUINTA DA LOMBA (BARREIRO)		471	8810334	MONTE REAL
69	8811386	TABOEIRA		270	8810884	ELVAS		472	8810450	PENALVA DO CASTELO
70	8811424	URGEZES		271	8810182	LORDELO (DOURO)		473	8810942	FUSETA
71	8810742	CALVÁRIO (LISBOA)		272	8810603	ERICEIRA		474	8810481	TAROUCA
72	8810812	ESTORIL		273	8810576	SANTA IRIA DE AZOIA		475	8810937	MONTE GORDO
73	8810171	PAÇOS DE FERREIRA		274	8810444	SEIA		476	8810016	VILA PRAIA DE ÂNCORA
74	8810327	MARINHA GRANDE		275	8810193	LAPA (PORTO)		477	8810924	ALMODÔVAR
75	8811329	S. JOSÉ (UISEU)		276	8810875	MONTEMOR-O-NOVO		478	8810736	AJUDA (LISBOA)
76	8811030	MACHICO		277	8810197	CARVALHIDO (PORTO)		479	8810943	MONCARAPACHO
77	8810771	ROMA (LISBOA)		278	8810339	PORTO DE MÔS		480	8810890	BORBA
78	8810609	CALDAS DA RAINHA		279	8811472	PORTO ALTO		481	8810987	MONCHIQUE
79	8810970	LAGOA (ALGARVE)		280	8810815	S. JOÃO DO ESTORIL		482	8811009	ANGRA DO HEROISMO
80	8811381	NORTE (MATOSINHOS)		281	8810839	PERO PINHEIRO		483	8810067	CABECEIRAS DE BASTO
81	8810213	DEVESAS (VILA NOVA DE GAIA)		282	8810225	VALADARES		484	8810536	SABUGAL
82	8810135	AEROPORTO DO PORTO		283	8811485	FILIPA DE LENCASTRE (BELAS)		485	8810077	MONCORVO
83	8810002	ALTO MINHO (VIANA DO CASTELO)		284	8810895	REGUENGOS DE MONSARAZ		486	8810458	CELORICO DA BEIRA
84	8810738	PRAÇA DO MUNICÍPIO (LISBOA)		285	8810802	BAIRRO NOVO (AMADORA)		487	8810545	ALCAINS
85	8810734	ALGES (LISBOA)		286	8810655	ENTRONCAMENTO		488	8810470	RESENDE
86	8810150	VILA DO CONDE		287	8810011	MONÇÃO		489	8810905	FERREIRA DO ALENTEJO

87	8810819	OEIRAS		288	8811376	S. PEDRO DA COVA (GONDOMAR)		490	8811044	CALHETA (MADEIRA)
88	8817031	LOURES SHOPPING		289	8810303	OLIVEIRA DO BAIRRO		491	8810711	LAVRADIO
89	8810825	SINTRA		290	8811476	BOLSA (LISBOA)		492	8810643	FERREIRA DO ZÉZERE
90	8810145	SANTO TIRSO		291	8811331	COLINA DO SOL		493	8810124	VILA POÚÇA DE AGUIAR
91	8811338	VALE DAS FLORES (COIMBRA)		292	8817042	PARQUE DAS NAÇÕES (LISBOA)		494	8810111	SANTA MARTA PENAGUIÃO
92	8810735	RESTELO (LISBOA)		293	8810149	VILA DAS AVES		495	8810894	REDONDO
93	8811337	ARRABIDA (VILA NOVA DE GAIA)		294	8810209	ASPRELA (PORTO)		496	8810596	ATOUGUIA DA BALEIA
94	8810379	PEDRULHA (COIMBRA)		295	8810575	BOBADELA (LOURES)		497	8810283	UNIVERSIDADE (AVEIRO)
95	8810775	S. JOÃO DE BRITO (LISBOA)		296	8810918	VILA NOVA DE SANTO ANDRÉ		498	8810314	PALHAÇA
96	8810196	FOZ DO DOURO (PORTO)		297	8810992	CALHETA (PONTA DELGADA)		499	8810337	MIRA DE AIRE
97	8810800	CARNAXIDE		298	8810665	ALMEIRIM		500	8810369	GOIS
98	8810833	RIO DE MOURO		299	8810146	CORONADO (SÃO ROMÃO)		501	8810988	VILA DE SAGRES
99	8811408	CANTARIAS (BRAGANÇA)		300	8810695	AZEITÃO		502	8810167	BAIÃO
100	8810791	TELHEIRAS (LISBOA)		301	8810614	ALENQUER		503	8810480	VILA NOVA DE PAIVA
101	8811041	NAZARÉ (FUNCHAL)		302	8811055	SANTO ANTONIO		504	8810084	MIRANDA DO DOURO
102	8810778	GRAÇA (LISBOA)		303	8810449	MANGUALDE		505	8810478	VOUZELA
103	8810585	VILA DE CANEÇAS		304	8810577	VIALONGA		506	8810237	PAÇOS DE BRANDÃO
104	8810758	LAPA (LISBOA)		305	8810518	VILAR FORMOSO		507	8810430	PENACOVA
105	8810245	OLIVEIRA DE AZEÍMIS		306	8810683	ALCOCHETE		508	8810872	ARRAIÓLOS
106	8810375	FERNAO DE MAGALHÃES (COIMBRA)		307	8810728	PRAGAL (ALMADA)		509	8810106	MURÇA
107	8811422	TAGUSPARK		308	8811048	RIBEIRA BRAVA		510	8811010	PRAIA DA VITÓRIA
108	8811349	CHARNECA DA CAPARICA		309	8810183	REBORDOSA (DOURO)		511	8810097	SABROSA
109	8810832	CACÉM		310	8810343	ANSIÃO		512	8810351	FIGUEIRO DOS VINHOS
110	8810045	RIBEIRÃO		311	8810733	BELEM (LISBOA)		513	8810069	CELORICO DE BASTO
111	8810342	POMBAL		312	8810835	BARCARENA		514	8810484	S. JOÃO DA PESQUEIRA
112	8810226	CANELAS (VILA NOVA DE GAIA)		313	8810963	SILVES		515	8810600	A.DOS-CUNHADOS
113	8811334	PRIOR VELHO		314	8811401	OLHOS DE ÁGUA		516	8810119	MONTALEGRE
114	8810221	AVINTES		315	8810282	VAGOS		517	8810493	SATÃO
115	8810152	PRAIA (PÓVOA DE VARZIM)		316	8810639	TORRES NOVAS		518	8810107	MESÃO FRIO
116	8810816	REBELVA (CARCAVELOS)		317	8810082	MACEDO DE CAVALEIROS		519	8810486	TABUAÇO
117	8810532	FUNDAO		318	8810659	ALFERRAREDE		520	8810497	MEDA
118	8810612	CASTANHEIRA DO RIBATEJO		319	8810278	GAFANHA DA NAZARÉ		521	8810951	AEROPORTO DE FARO
119	8810638	TOMAR		320	8810929	ODEMIRA		522	8810100	ALLUO
120	8811330	GALIZA (PORTO)		321	8810673	BENAVENTE		523	8811005	RIBEIRA GRANDE
121	8810330	LEIRIA		322	8811493	CÔRTE REAL (TERCEIRA)		524	8810362	PENELA
122	8810961	AREIAS DE S. JOÃO		323	8810594	PONTINHA (LISBOA)		525	8810354	ALVAÍZERE
123	8810284	ALBERGARIA-A-VELHA		324	8810572	SACAVÉM		526	8810669	GOLEGÃ

124	8810713	QUINTA DO CONDE		325	8810222	OLIVEIRA DO DOURO (VILA NOVA DE GAIA)		527	8810664	ALPIARÇA
125	8810170	PAREDES		326	8810718	COVA DA PIEDADE (ALMADA)		528	8810118	RIBEIRA DE PENA
126	8810094	MIRANDELA		327	8811006	VASCO DA GAMA (PONTA DELGADA)		529	8810667	CHAMUSCA
127	8810808	AMADORA		328	8810153	CAXINAS (VILA DO CONDE)		530	8811008	AEROPORTO DE PONTA DELGADA
128	8810799	LINDA-A-VELHA		329	8810586	POVOA DE SANTO ADRIÃO		531	8811023	MADALENA (PICO)
129	8810694	PINHAL NOVO		330	8810214	CARVALHOS		532	8810452	FORNOS DE ALGODRES
130	8810805	DAMAIA		331	8810338	MACEIRA LIZ		533	8810866	AVIS
131	8810141	SENHORA DA HORA		332	8810869	ÉVORA		534	8810526	BELMONTE
132	8810230	ESPINHO		333	8810316	MEALHADA		535	8810887	SOUSEL
133	8811406	SANTA TECLA (BRAGA)		334	8811045	CÂMARA DE LOBOS		536	8810657	VILA NOVA DA BARQUINHA
134	8810376	S. JOSE (COIMBRA)		335	8810289	BRANCA		537	8811050	S. VICENTE (MADEIRA)
135	8810270	AVEIRO		336	8810043	LOUSADO		538	8810952	ESTOI
136	8810790	CAMPO GRANDE (LISBOA)		337	8810692	ALHOS VEDROS		539	8810495	AGUIAR DA BEIRA
137	8810687	BONFIM (SETÚBAL)		338	8810935	VILA REAL DE SANTO ANTONIO		540	8810073	VILA FLOR
138	8810050	FAFE		339	8817045	FERNÃO FERRO		541	8810014	PAREDES DE COURA
139	8810154	PENAFIEL		340	8811499	COIMBRÕES (VISEU)		542	8810333	CARANGUEJEIRA
140	8810966	ARMAÇÃO DE PERA		341	8810574	CAMARATE		543	8810851	NISA
141	8810817	PAEDE		342	8810724	CORROIOS (ALMADA)		544	8810668	CONSTÂNCIA
142	8810397	CANTANHEDE		343	8810971	PRAIA DO CARVOEIRO (ALGARVE)		545	8810936	CASTRO MARIM
143	8810803	ALFRAGIDE		344	8810920	SINES		546	8810882	VIANA DO ALENTEJO
144	8811380	LOJA DO CIDADÃO (LISBOA)		345	8810258	ESMORIZ		547	8810494	SERNANCELHE
145	8810132	S. MAMEDE DE INFESTA		346	8817038	DOLCE VITA TEJO		548	8810510	FIGUEIRA DE CASTELO RODRIGO
146	8810693	MOITA		347	8810021	BARROSELAS		549	8811013	AEROPORTO DAS LAJES (TERCEIRA)
147	8810130	MATOSINHOS		348	8810223	AROUCÁ		550	8810705	SESIMBRA
148	8810764	CAMPOLIDE (LISBOA)		349	8810578	S. JOÃO DA TALHA		551	8810893	MOURÃO
149	8810240	SANTA MARIA DA FEIRA		350	8810053	CALDAS DE VIZELA		552	8810543	PENAMACOR
150	8811035	MONUMENTAL (FUNCHAL)		351	8811423	AVENIDA (LOULÉ)		553	8810648	RIACHOS
151	8811490	FEITEIRA (CARVALHOS)		352	8810975	ALVOR		554	8810925	OURIQUE
152	8811397	MAR E GUERRA (FARO)		353	8810015	VILA NOVA DE CERVEIRA		555	8811014	VELAS
153	8811405	ZONA INDUSTRIAL (CASTELO BRANCO)		354	8810057	FEIRA NOVA (AMARES)		556	8810482	ARMAMAR
154	8810095	PESO RÉGUA		355	8810503	S. MIGUEL (GUARDA)		557	8810880	PORTEL
155	8810064	VILA DE PRADO		356	8811028	AEROPORTO DO FUNCHAL		558	8811003	VILA FRANCA DO CAMPO
156	8810419	OLIVEIRA DO HOSPITAL		357	8810162	TERMAS DE S. VICENTE		559	8810260	MURTOSA
157	8811336	GAIA		358	8810302	ANADIA		560	8810091	VIMIOSO
158	8811039	CANIÇO (SANTA CRUZ MADEIRA)		359	8810676	SAMORA CORREIA		561	8810653	MAÇÃO

159	8810955	QUARTEIRA		360	8810900	GRÂNDOLA		562	8810438	CANAS DE SENHORIM
160	8810584	ODIVELAS		361	8810672	CORUCHE		563	8810008	MELGAÇO
161	8810763	CORTES (LISBOA)		362	8810025	ARCOS DE VALDEVEZ		564	8810521	MANTEIGAS
162	8810682	S. PEDRO (MONTUJO)		363	8810611	VILA FRANCA DE XIRA		565	8810877	MORA
163	8810626	RIO MAIOR		364	8810115	CHAVES		566	8810886	MONFORTE
164	8810750	D. AFONSO HENRIQUES (LISBOA)		365	8810681	MONTUJO		567	8810860	FRONTEIRA
165	8811394	PENHA (FARO)		366	8810640	OURÉM		568	8810856	CASTELO DE VIDE
166	8810179	FREAMUNDE		367	8810591	VENDA DO PINHEIRO		569	8811021	LAJES DO PICO
167	8810891	ESTREMOZ		368	8810126	VALPAÇOS		570	8810085	FREIXO ESPADA À CINTA
168	8810188	BATALHA (PORTO)		369	8810266	ESGUEIRA (AVEIRO)		571	8810061	TERRAS DE BOURO
169	8810003	PONTE DE LIMA		370	8810382	S. MARTINHO DO BISPO		572	8810840	VILA (SINTRA)
170	8810065	VILA VERDE		371	8810897	VILA VIÇOSA		573	8810116	MONDIM DE BASTO
171	8810721	LARANJEIRO (ALMADA)		372	8810388	MONTEMOR-O-VELHO		574	8810120	BOTICAS
172	8810814	S. PEDRO DO ESTORIL		373	8810881	VENDAS NOVAS		575	8810885	ARRONCHES
173	8810163	LIXA		374	8810400	FEBRES		576	8810997	S. ROQUE (PONTA DELGADA)
174	8810140	LEÇA DA PALMEIRA		375	8810359	CONDEIXA		577	8810490	PENEDONO
175	8810522	COVILHÃ		376	8810004	CAMINHA		578	8811051	ARCO DA CALHETA
176	8810500	GUARDA		377	8810617	CARREGADO		579	8810854	ALTER DO CHÃO
177	8810202	GARCIA DE ORTA (PORTO)		378	8810691	BAIXA DA BANHEIRA		580	8810562	VILA VELHA RODÃO
178	8810229	PRAIA DA GRANJA		379	8810837	MIGUEL BOMBARDA (QUELUZ)		581	8811000	POVOAÇÃO
179	8810589	HERÓIS DE CHAMITE (ODIVELAS)		380	8810836	BELAS		582	8811002	VILA DO PORTO
180	8810255	OVAR		381	8810981	LUZ (LAGOS)		583	8810858	CRATO
181	8810136	CASTELO DA MAIA		382	8810972	S. BARTOLOMEU DE MESSINES		584	8810551	OLEIROS
182	8811348	LAMEIRAS (GUIMARÃES)		383	8810615	CADAVAL		585	8811025	SANTA CRUZ DAS FLORES
183	8811498	MIRAFLORES		384	8810416	MORTÁGUA		586	8810999	NORDESTE
184	8810040	ESPOSENDE		385	8810571	PÓVOA DE SANTA IRIA		587	8810760	ASSEMBLEIA DA REPÚBLICA (LISBOA)
185	8810719	PRAÇA (ALMADA)		386	8810579	FORTE DA CASA		588	8810598	ÓBIDOS
186	8810762	CAMPO DE OURIQUE (LISBOA)		387	8810134	ÁGUAS SANTAS (MAIA)		589	8810506	ALMEIDA
187	8810784	OLIVAIS SUL (LISBOA)		388	8810732	JUNQUEIRA (LISBOA)		590	8810559	VILA DE REI
188	8810688	N. S. DA ANUNCIADA (SETÚBAL)		389	8810242	RIO MEÃO		591	8810889	ALANDROAL
189	8811004	ANTERO DE QUENTAL (PONTA DELGADA)		390	8810597	BOMBARRAL		592	8810994	RABO DE PEIXE
190	8810982	PORTAS DE PORTUGAL (LAGOS)		391	8810176	AREOSA		593	8810373	PAMPILHOSA DA SERRA
191	8810332	BATALHA		392	8810296	AGUADA DE CIMA		594	8811047	PORTO MONIZ
192	8811450	ARAUCÁRIA		393	8810590	MALVEIRA		595	8810670	SARDOAL
193	8810796	SETE RIOS (LISBOA)		394	8810865	PONTE DE SÔR		596	8811015	CALHETA (S. JORGE)
194	8810129	PARQUE (MATOSINHOS)		395	8810396	FIGUEIRA DA FOZ		597	8810944	ALCOUTIM

195	8810208	CONDE DE FERREIRA (PORTO)		396	8810026	PONTE DA BARCA		598	8811024	LAJES DAS FLORES
196	8810156	MARCO DE CANAVEZES		397	8810410	TOCHA		599	8810902	ALVITO
197	8817034	ALFRAGIDE SHOPPING		398	8810365	ARGANIL		600	8811022	S. ROQUE DO PICO
198	8810363	LOUSÃ		399	8810472	OLIVEIRA DE FRADES		601	8810349	CASTANHEIRA DE PERA
199	8810464	LAMEGO		400	8810263	AVANCA		602	8810907	BARRANCOS
200	8811339	MARQUÊS DE POMBAL (LEIRIA)		401	8810408	MIRA		603	8811026	CORVO
201	8810587	LOURES		402	8810556	SERTÃO		604	8810862	MARVÃO
				403	8810904	CUBA (ALENTEJO)		605	8810752	SANTA MARTA (LISBOA)
								606	8810186	MUNICÍPIO (PORTO)

8.3. Anexo 3

Como não é possível visualizar na matriz U todos os códigos de cada posição, em baixo está a lista com todos os códigos de cada posição e os que são visíveis identificados a amarelo, 1º nível.

1º Nível	2º Nível	3º Nível	4º Nível	5º Nível	6º Nível	7º Nível	1º Nível	2º Nível	3º Nível	4º Nível	5º Nível	6º Nível	7º Nível	8º Nível
1	606						151	178						
2	26						152	259	290					
3							153							
4	21						154							
5							155	227	233					
6							156							
7	27						159							
8	16						160							
9	50						161							
10	30						163							
11							164							
12	52						166							
13	78						169	265						
14	111						170							
15							172							
17	48						174	194	243					
18	49						175	180						
19	29						176							
20							177							
22							179							
23							181	183						
24	34	86					182							
25	35	38					184							
28							185	364						
31	89	97					186							
32							188	306	385					
33	108						189							
36							190							
37							192							
39	121						193	201						
40	103						195	240						
41	73						197							
42							198							
43	44						200							
45							202	283	387	396	448			
46							204							
47	114	237					205							
51	64	106					206							
53	168	203					207	302						
54							208	277	358					
55							211							
56	69	70					212	278	301					
57							213	363						
58							214							
59							216							

60	84						220							
61							221							
62	77						222	435						
63	105	139	224				223							
65	234						228							
66	120						230	275						
67	238						231							
68							235	239	286					
71							236	248						
72							242							
74							244							
75	171						245	268						
76	281	362					246	264	307	605				
79	98						247	366						
80	93	157	260				250	270	338					
81	83	144	145				251	407	466					
82							252	391	426					
85	91	134					253	313	408					
87	327						254	295						
88							255	372						
90							256	350						
92	143						257	321	370					
94	162						258	291						
95							261	300						
96	232						262	324						
99							266	353						
100							267	346	405					
101	158						269	317						
102	218	241					271	394						
104							272							
107							273	458						
109	133						274							
110	167						276	418						
112	122						279	305	336					
113							280	308						
115							282							
116	215						284	296	304	359				
117	217						285	323	379					
118	288						287							
119							289	315						
123	191	219	229				292	293	347	431				
124							294							
125	196						297							
126							298							
127	128	147					299							
129	165						303							
130							309	417						
131							310	402	410					
132							311							
135	137						312							
136	249						314	380						
138	226						316							
140	263						318	329	349					
141	187	225					319							
142	199	209					320							
146							322	454	523					
148	173						325							
149	210						326							
150	334						328							
331	337	341	373	389	492		330							
332							427							
333							433	471						
335	374						437	495	518	548				
339							438							
340	390						439							
342							443	531						
343	388	447					446	452						
344							450							
345							453	494	554	560	561			
348	368						455							
351							456	550						
352	354	432					457	481	529	569				
355							460	477	479					
356	497	530	542	588	594	596	461	506						
357	442						462							
360							464	559						
361							465	475	540					
365							468							
367	386						469							
369	375						470	493						

371							472	503	528					
376							478							
377							480	496	508	538				
378							482							
381	392	429	474	526			484							
382	423						485	486						
383	409						488	489	541					
384	483						490	491	514	534				
393							500	584						
395							501	521						
397	434	445	504				507	513						
398	413	440	451				509	532	575					
399	428	463					510	558						
400	436	487					511	543						
401							515							
403	553						519	533	551	565	579			
404	473						520	552	574					
406	476						522							
411							524	525	536	562	573			
412							527	539						
414							535	547	556	568	589			
415	459	499					544	564	586	595				
416	498	502	512				545							
419							546	557						
420	505	516					549	555	582					
421	449	467	537	581			563							
422							566	580	602					
424	430	441	444	517			567	593	598					
425							570	592						
							571	578	590	601				
							572							
							576	585	600					
							577	583	587	591	597	599	603	604

